

ASSESSING EMERGENT BILINGUALS:  
TEACHER KNOWLEDGE AND READING INSTRUCTIONAL PRACTICES

by

LAURA ASCENZI-MORENO

A dissertation submitted to the Graduate Faculty in Urban Education in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2012

© 2012

LAURA ASCENZI-MORENO

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Urban Education in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Ofelia García\_\_\_\_\_

\_\_\_\_\_  
Date

\_\_\_\_\_  
Chair of Examining Committee

Anthony Picciano\_\_\_\_\_

\_\_\_\_\_  
Date

\_\_\_\_\_  
Chair of the Examining Committee

Ofelia García\_\_\_\_\_

Kate Menken\_\_\_\_\_

Anthony Picciano\_\_\_\_\_  
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

## Abstract

### ASSESSING EMERGENT BILINGUALS: TEACHER KNOWLEDGE AND READING INSTRUCTIONAL PRACTICES

by

Laura Ascenzi-Moreno

Adviser: Ofelia García

Assessments are viewed as primary vehicles for improving the educational outcomes of all students since they can lay the foundation for effective teaching practices. However, assessment can only achieve this effect of supporting student learning if the knowledge that teachers gain from using them is put into direct use in classrooms. This process of administering assessments, analyzing them, learning from the results, and subsequently tailoring instruction based on what has been learned about students is referred to as the *assessment-instructional cycle*.

The assessment-instructional cycle is critical for all students. Yet, assessments of emergent bilingual students, in this case those students who are becoming bilingual by developing English language and literacy (often referred to as English Language Learners or Limited English Proficient students), most often do not accurately capture these students' knowledge. The problem lies in that for these students, assessments in English measure content knowledge as well as *language* (Abedi, 2009; García 2009). This obscures teachers' understanding of what these students know and may steer the assessment-instructional cycle off course.

Using interviews and surveys, this mixed-methods study focuses on how teachers of emergent bilinguals view and use summative and formative assessment. The study also attempts

to ascertain the kinds of knowledge that these teachers gain from the use of assessments, as well as the consequences that acquired knowledge has on their practices to teach reading.

In the climate of test-based accountability, teachers are caught in a cycle of *ritualized assessment practices*. Ritualized assessment practices direct teachers to sort and group students under the guise of “analysis,” and do not engage teachers in a solid examination of bilingual students’ reading development. Ritualized assessment practices ultimately do not yield teacher knowledge that is meaningful to instruction. Furthermore, these ritualized practices change the character and use of these assessments – summative assessments are used formatively and formative assessments become summative. This study provides evidence that both summative and formative assessments are missed opportunities for teacher learning and do not fulfill the potential of providing teachers with a solid knowledge base of their bilingual students’ reading development so as to meaningfully direct instructional practices.

## **Acknowledgements**

The journey to completing this project has been long and I have been lucky to be accompanied by many people. I would like to whole-heartedly thank my “team” who has inspired and motivated me throughout.

I would like to start with my committee members. Ofelia García has been my constant and unwavering advocate. She has set firm expectations, offered generous support, and kind words throughout the process. Much of my professional identity and research interests are the result of her teaching and guidance. Kate Menken has been an invaluable mentor to me. She has shown me the “research ropes” and I am grateful for all the times she listened to my ideas and offered me support – professional and otherwise. I would also like to thank Martin Ruck who gave me meticulous comments on the dissertation proposal. Lastly, I would like to acknowledge Anthony Picciano who has been generous with his time and provided critical feedback to help move this project forward.

In addition friends and colleagues have been essential to completing this dissertation. In particular, I want to thank Berky Lugo-Salcedo and Sara Lippi who gave their professional input throughout. Their contributions for creating and editing the list of instructional practices included in the survey were essential. Also, I would like to acknowledge Inés Ellis-Guardiola for her help in field testing the survey. Meral Kaya has been my “angel” at Brooklyn College and without her assistance there, I would have not been able to concentrate as much as I have on this work. Many of the ideas for this dissertation were incubated with other Graduate Center colleagues including Nelson Flores, Zoila Tazi, Kirsten Cole, Ivana Espinet, Liza Pappas, and Haiwen Chu. Thank you all.

I am also indebted to all the teachers and administrators who opened up their schools to this work and spoke about a subject that is charged both politically and emotionally. Without their help, this work would not have been possible. It is my sincere hope that this work illuminates ways in which assessments can be used for the benefit of emergent bilingual students and contributes to teacher knowledge.

Finally I would like to thank the family portion of my “team.” I am thankful for the help my in-laws, Liliana Cevoli and Gian Carlo Clementi, provided while hosting us in their home which allowed me to write during the summer. My mother’s constant presence and help was essential to me. It would have been truly impossible without her support to complete this dissertation. I also want to recognize my son, Matteo, who has enriched and deepened the happiness of my life in inconceivable ways. Finally, I would like to recognize my husband, Gian Luca, who supported me through thick and thin and through big and small ways throughout this journey.

## Table of Contents

	<b>Page</b>
<b>Abstract</b> .....	<i>iv</i>
<b>Acknowledgements</b> .....	<i>vi</i>
<b>List of Tables</b> .....	<i>xv</i>
<b>List of Figures</b> .....	<i>xvii</i>

### **Chapter 1:**

<b>Introduction</b> .....	1
Research Questions .....	4
Purpose: The Weaving of Personal and Academic Trajectories .....	4
Accountability: The Context for Assessment.....	6
New York City’s Emergent Bilinguals: A Critical Piece in New York City Accountability Reform .....	8

### **Chapter 2:**

<b>Literature Review</b> .....	15
Part 1: The Evolution of Teacher Knowledge .....	15
Part 2: Assessment & Accountability .....	24
Assessment .....	28
Accountability .....	28
Part 3: Assessment of Emergent Bilinguals .....	33
Bilingualism & Assessment of Content .....	33
Summative Assessments & Bilingualism .....	38
Accommodations .....	42
Test Translations .....	44
Parallel Tests – Transadaptations .....	45

	<b>Page</b>
Conceptual Scoring .....	45
Summative Assessments: Problems & Directions .....	46
Formative Assessment & Bilingualism .....	49
Part 4: Reading Development of Emergent Bilinguals .....	55
Summary .....	61
<b>Chapter 3:</b>	
<b>Methodology</b> .....	<b>63</b>
Introduction: A Mixed Methods Study .....	63
Part 1: Research Design .....	64
Part 2: Instrumentation .....	68
Timeline .....	69
Interviews .....	70
Surveys .....	72
Part 3: Assessments in Focus .....	81
English Language Arts Exam (ELA) .....	82
New York State English as a Second Language Achievement Test (NYSESLAT).....	83
Running Records .....	84
Reading Conferences .....	86
Part 4: Sampling, Participants & Research Sites .....	87
Interviews .....	87
School 1 .....	89
School 2 .....	90

	<b>Page</b>
School 3 .....	91
School 4 .....	92
Survey .....	95
Part 5: Data Analysis .....	97
Interviews .....	97
Survey .....	98
Part 6: Limitations .....	100
Interviews .....	100
Survey .....	101
Summary .....	103

#### **Chapter 4:**

<b>Results from the Interviews</b> .....	<b>104</b>
Part 1: Teachers Speak about the English Language Arts Exam (ELA), a Summative Assessment .....	105
A Lack of Trust: Critiques of the Use of ELA Data .....	108
Operating Instructions: Knowledge and the ELA Exam .....	115
Mining the ELA for Instruction: “We are using what you can use.” .....	121
Involving Students: An Opportunity to Rescue ELA Data .....	129
Learning and Not Learning from the ELA: A Summary .....	130
Part 2: Teachers Speak about Running Records, a Formative Assessment .....	132
Seeing Students as Readers: Student Knowledge Acquisition from Running Records .....	135

	<b>Page</b>
“I do what I am told to do.” Going through the Motions in Administering Running Records .....	139
Owning the Assessment: The Acquisition of Assessment Knowledge .....	147
Seeing the Whole Child as a Reader: The Acquisition of Bilingual Knowledge .....	151
Assessing Students through Spanish Running Records: Inequities from the Start .....	156
Language Transfer: Bilingualism Takes Care of Itself .....	160
Pedagogical Content of Reading: A Singular Voice .....	162
Instructional Practices: Assessment Behavior Disconnected from Data .....	166
Grouping .....	166
Mini-Lesson Design .....	168
Vocabulary .....	169
Feedback .....	171
Learning and Not Learning from Running Records: A Summary .....	173
General Summary .....	175

### **Chapter 5:**

<b>Findings from the Surveys</b> .....	<b>178</b>
How frequently are ELA and Running Record Data Used? .....	180
What do Teachers Learn from the Assessment Data of Emergent Bilinguals .....	185
Pedagogical Content Knowledge .....	187
Student Knowledge .....	189
Assessment Knowledge .....	190

	<b>Page</b>
Bilingual Knowledge .....	192
Type of Knowledge Gained .....	193
Is How Often an Assessment used Related to Greater Levels of Acquisition of Teacher Knowledge? .....	195
Do Teachers Report Changes in Instructional Practices as a Result of Examining Data? .....	197
Instructional Category: Grouping .....	198
Instructional Category: Reading Strategies .....	200
Prediction .....	201
Questioning .....	202
Summarizing .....	203
Analysis of Text .....	204
Summary: Reading Strategies Category .....	205
Instructional Category: Word Attack Skills .....	205
Sight Words .....	205
Decoding .....	206
Summary: Word Attack Instructional Category .....	207
Instructional Category: Vocabulary .....	208
Texts with Limited Vocabulary .....	208
Teaching Students Required Vocabulary .....	209
Teach Meaning of Unknown Words .....	210
Summary: Vocabulary Instructional Category .....	211
Instructional Category: Background Knowledge .....	212

	<b>Page</b>
Activating Prior Knowledge & Providing Students with Background Knowledge .....	213
Summary: Background Knowledge Category .....	215
Instructional Category: Use of Home Language .....	215
Summary: Use of Home Language .....	217
Instructional Category: Oral Language .....	217
Asking Students to Speak about Texts & Engaging Students in Partner Talk .....	218
Summary: Oral Language .....	220
Instructional Category: Assessment .....	220
Providing Students with Opportunities to Practice Assessment .....	221
Observing Students as an Instructional Practice .....	222
Summary of Assessment Instructional Practices .....	223
Summary of Analysis of Instructional Practices .....	224
How are Changes in Instructional Practices Associated with Teacher Knowledge? .....	228
Summary .....	235

## **Chapter 6:**

<b>Discussion &amp; Conclusion</b> .....	237
General Finding 1 .....	239
General Finding 2 .....	242
General Finding 3 .....	245
General Finding 4 .....	248
General Finding 5 .....	250

	<b>Page</b>
Conclusion & Implications .....	252
Recommendations .....	254
<b>Appendix 1.</b> Qualitative Research Questions .....	257
<b>Appendix 2.</b> Survey .....	261
<b>Appendix 3.</b> Description of Quantitative Variables .....	276
<b>Appendix 4.</b> Sample of New York State English Language Arts Exam 2009, 4 <sup>th</sup> Grade .....	282
<b>Appendix 5.</b> Running Record Sample .....	284
<b>Appendix 6.</b> Types of Teachers that Work with Emergent Bilingual Students .....	286
<b>Appendix 7.</b> Codes for Qualitative Analysis .....	287
<b>Appendix 8.</b> Example of Description of Fountas & Pinnell Reading Levels .....	290
<b>Appendix 9.</b> Text Level Correlation Chart .....	291
<b>References</b> .....	292

## List of Tables

<b>Table</b>	<b>Page</b>
1. Shulman’s Model of Teacher Knowledge .....	17
2. Danielson’s Framework for Teaching: Focus on Teacher Knowledge .....	19
3. Heritage’s Types of Teacher Knowledge Required for Assessment .....	20
4. Ascenzi-Moreno’s Model for Knowledge for Teachers of Emergent Bilinguals .....	23
5. Frequent Types of Accommodations Implemented for Emergent Bilinguals .....	44
6. Research Question & Methodology Matrix .....	65
7. Overview of Variables used in Quantitative Analysis .....	76-7
8. Instructional Practices by Categories .....	79
9. Demographics and Socio-economic Characteristics of Schools .....	88
10. Descriptive Statistics for Sample of Teachers Interviewed and Teachers Surveyed ..	94
11. Percent Use of ELA Assessment .....	182
12. Percent Use of Running Records .....	183
13. Frequency Distribution for Pedagogical Content Knowledge .....	188
14. Frequency Distribution for Student Knowledge .....	190
15. Frequency Distribution for Assessment Knowledge .....	191
16. Frequency Distribution for Bilingual Knowledge .....	192
17. Comparison of Frequency Distribution for Types of Teacher Knowledge .....	194
18. Correlations between Assessment Use & Types of Teacher Knowledge .....	196
19. Frequency Distribution for Grouping as an Instructional Practice .....	199
20. Frequency Distribution for Prediction as an Instructional Practice .....	201
21. Frequency Distribution for Questioning as an Instructional Practice .....	202
22. Frequency Distribution for Summarizing as an Instructional Practice .....	203

<b>Table</b>	<b>Page</b>
23. Frequency Distribution for Analysis of Texts as an Instructional Practice .....	204
24. Frequency Distribution for Sight Words as an Instructional Practice .....	206
25. Frequency Distribution for Decoding as an Instructional Practice .....	207
26. Frequency Distribution for Texts with Limited Vocabulary as an Instructional Practice .....	208
27. Frequency Distribution for Teach Students Required Vocabulary as an Instructional Practice .....	210
28. Frequency Distribution for Teach Meaning of Unknown Words as an Instructional Practice .....	211
29. Frequency Distribution for Activate Prior Knowledge as an Instructional Practice ..	213
30. Frequency Distribution for Provide Students with Background Knowledge as an Instructional Practice .....	214
31. Frequency Distribution for Provide Students with Texts in Native Language as an Instructional Practice .....	216
32. Frequency Distribution for Ask Students to Speak about Texts as an Instructional Practice .....	218
33. Frequency Distribution for Engage Students in Partner Talk as an Instructional Practice .....	219
34. Frequency Distribution for Provide Students with Opportunities to Practice Assessment as an Instructional Practice .....	221
35. Frequency Distribution for Observe Students as an Instructional Practice .....	223
36. Comparison of Frequency Distribution for Types of Instructional Practices .....	224-5
37. Correlations between Teacher Knowledge and Instructional Practices (ELA) .....	230
38. Correlations between Teacher Knowledge and Instructional Practices (Running Records) .....	233
39. Ascenzi-Moreno’s Revised Model of Knowledge for Bilingual Teachers .....	247

## List of Figures

<b>Figure</b>	<b>Page</b>
1. CDF of ELA & Running Record Pedagogical Content Knowledge .....	189
2. CDF of ELA & Running Record Student Knowledge .....	190
3. CDF of ELA & Running Record Assessment Knowledge .....	191
4. CDF of ELA & Running Record Bilingual Knowledge .....	192
5. CDF of Grouping as an Instructional Practice .....	199
6. CDF of Prediction as an Instructional Practice .....	201
7. CDF of Questioning as an Instructional Practice .....	202
8. CDF of Summarizing as an Instructional Practice .....	203
9. CDF of Analysis of Texts as an Instructional Practice .....	204
10. CDF of Teach Sight Words as an Instructional Practice .....	206
11. CDF of Decoding as an Instructional Practice .....	207
12. CDF of Texts with Limited Vocabulary as an Instructional Practice .....	209
13. CDF of Teach Students Required Vocabulary as an Instructional Practice .....	210
14. CDF of Teach Meaning of Unknown Words as an Instructional Practice .....	211
15. CDF of Activate Prior Knowledge as an Instructional Practice .....	214
16. CDF of Provide Students with Background Knowledge as an Instructional Practice .....	215
17. CDF of Provide Students with Texts in Native Language as an Instructional Practice .....	217
18. CDF of Ask Students to Speak about Texts as an Instructional Practice .....	219

<b>Figure</b>	<b>Page</b>
19. CDF of Engage Students in Partner Talk as an Instructional Practice .....	220
20. CDF of Provide Students with Opportunities to Practice Assessment as an Instructional Practice .....	222
21. CDF of Observe Students as an Instructional Practice .....	223

## CHAPTER 1:

### Introduction

In recent decades, assessments have been viewed by policy makers as a main catalyst in educational reform; they are charged with improving education for all students (Darling-Hammond & Falk, 1997; Ladd, 2008; Linn, 2003). In order for this vision to be realized, assessments must provide the foundation for effective teaching practices (Heritage, Kim, Vendlinski & Herman, 2009). Yet, there are several challenges to the intended goal of achieving educational equity through assessment, starting with the problem that not all students have the same assessment needs (Abedi, Hofstetter & Lord, 2004; Solano-Flores, 2008). Another challenge, and the one which is the focus of this research, is that teachers must have the ability to interpret assessment data and use it to develop effective instruction for the specific students they are working with (Heritage, 2007).

The assessment needs of emergent bilingual students<sup>1</sup> differ from those of the general population. The assessment of these students is complicated by a range of factors starting with the ability of assessments to capture with precision these students' knowledge rather than

---

<sup>1</sup> The term, 'emergent bilingual' developed by Garcia (2009) is preferred and will be used rather than 'English Language Learner' in this dissertation. Although both terms refer to similar populations of students, they indicate slightly different groups and emerge from different perspectives on the acquisition of language. English Language Learner (ELL), the term officially used by federal, state, and city agencies refers to students whose home language is not English and who are in the process of acquiring academic English. This definition highlights students' lack of English proficiency. In contrast, the term 'emergent bilingual' describes students who are in the *process* of learning *two or more* languages. It calls attention to two key points about bilingual development in general. First, bilingualism exists within a continuum, rather than as fixed categories. Secondly, as emergent bilinguals acquire English, they are developing bilingually. Therefore, the term 'emergent bilingual,' stems from the perspective of valuing a child's knowledge of and in two languages. Emergent bilinguals can also include students whose home language is English and are acquiring Spanish and will be used as such when referring to this particular population. The term 'emergent bilingual' will be used throughout this thesis to refer to the students who are the focus of this study. The term English Language Learner will be used when referring to studies or legislation which use this term.

confounding their performance with language ability (Abedi, 2009; García, 2009; GAO 2006; Keiffer, Lesaux & Snow, 2008).<sup>2</sup> The difficulty of accurately assessing bilinguals may obscure educators' understanding of what these students know and, as a result, may interrupt the desired flow from assessment to appropriate instruction for these students.

While reform is founded on the ability of assessments to spur instructional practices that directly improve student learning, it is unclear how teachers of emergent bilingual students make sense of the data from assessments these students take. Improvements to student learning can only result from assessment data, if the *assessment-instructional cycle* – the process of administering assessments, analyzing them and tailoring instruction for students – adequately suits the characteristics and needs of the intended student population. The knowledge that teachers generate from assessments of emergent bilinguals and how they apply that knowledge to improve instruction for these students is pivotal to ensuring that the assessment-instructional cycle indeed functions for this population of students. For both of these reasons – the ability of assessments to accurately capture emergent bilingual students' learning and the uncertain knowledge base from which teachers work to interpret data about these students – the promise of assessment to improve emergent bilingual students' learning is far from assured and has not been fulfilled.

This doctoral dissertation examines if and how teacher knowledge gained from the assessment of emergent bilinguals mediates the implementation of reading instructional practices for emergent bilingual students. These two factors – teacher knowledge and shifts in

---

<sup>2</sup> Emergent bilinguals are merged into the general pool of students who take mandated, large-scale standardized tests one year after arrival to the US.

instructional practice that result from teacher knowledge – are important ways in which assessment can be used to directly impact emergent bilingual students’ learning.

Specifically, this research examines what knowledge is generated by teachers from the assessment data of emergent bilinguals from two currently favored types of literacy assessments – large-scale standardized and classroom-based assessments – and what changes in instructional practices result from changes in teacher knowledge resulting from analysis of this data.<sup>3</sup> This research acknowledges the problematic nature of many assessments administered to emergent bilingual students. Rather than seeking to change assessments per se, this research aims to understand what knowledge teachers gain from assessments, if any, and if this acquired knowledge impacts instructional practices. Specifically, this study will be limited to the use of reading assessment data by elementary school (grades 3-5) teachers of emergent bilinguals. This research focuses on these grades because although the effect of testing has been documented with high school students, the effect of assessment on students in elementary schools has not studied to the same extent. This study intends to complement research that has illustrated the effects of testing on emergent bilingual students in high schools by providing an additional perspective on teacher knowledge, elementary school students, and formative assessments (Menken, 2008).

---

<sup>3</sup> These two types of assessments will be further defined in the review. Large-scale standardized assessments will later be referred to as *summative* assessments. For this study, classroom-based assessments will refer to “formal” formative assessments such as running records that may be part of a pre-packaged assessment kit. This type of assessment is *formative*.

## Research Questions

The study is guided by the following research questions:

1. What kinds of knowledge do teachers<sup>4</sup> of emergent bilinguals gain from examination of reading assessment data of their students?
  - How does the knowledge that these teachers gain differ when using large-scale standardized assessments and using classroom-based assessments?
2. What changes, if any, in reading instructional practices for emergent bilingual students result from teachers of emergent bilinguals' *use* of large-scale standardized assessments and/or classroom-based assessments?
3. What changes, if any, in reading instructional practices for emergent bilingual students result from *the types of knowledge* gained by teachers of emergent bilinguals from large-scale standardized assessments and/or classroom-based assessments?

## Purpose: The Weaving of Personal and Academic Trajectories

The purpose of this research is to provide insight into how teacher knowledge develops from the assessment of emergent bilinguals and if teacher knowledge acts as a bridge between assessment and instructional practices for these students. The impetus for this research stems from the intersection of both personal and academic interests.

As a dual language bilingual teacher in New York City schools for over a decade and later as an instructional coach, I became very familiar with the “assessment agenda” in schools. Assessments, both large-scale standardized and classroom-based, are deemed to improve instruction, by providing us, teachers, with information about our students. Data-driven instruction has become a powerful influence on how school administrators and teachers conduct

---

<sup>4</sup> Two sets of teachers of emergent bilinguals were part of this study. Teachers interviewed were entirely comprised of dual language bilingual teachers. Teachers who completed the survey were ESL, transitional bilingual, and dual language bilingual teachers.

their practice. As a teacher and instructional coach, I attended numerous professional development sessions about how to use data. As a coach, I spent countless hours pouring over school-wide reading data about our students. Although I found some of this time with data helpful in gaining a sense of trends of our students' reading, the question that I kept coming back to was: what were teachers of emergent bilinguals learning about their students' reading from working with data? Furthermore, was the knowledge that these teachers gained "useable" in the sense that it could be directly applied by teachers to create instructional practices suited to emergent bilinguals' learning needs?

As a graduate student, these questions reemerged. Research about assessment of emergent bilinguals indicates that large-scale standardized assessments often offer educators an indistinct picture of these students' content-related skills because language is confounded with content. On the other hand, while formative assessments have been found to impact student learning, few studies have been conducted to verify if this positive flow between formative assessment and gains in learning holds true for emergent bilingual students.

The merging of these two areas of my life – my work in schools and my work as a graduate student – have led me to pursue the question of how and what kinds of knowledge teachers obtain from the process of analyzing data of emergent bilingual students and if this knowledge acquired from data affects their teaching of these students. This question is timely given that "data-driven" decision-making and instruction is heavily funded federally and impacts school curricular priorities. Although accountability will be discussed in depth in the literature review, a brief overview of issues related to accountability is presented here to give context to this study. Furthermore, I also offer information on New York City's emergent bilingual

students and the programs that serve them. These topics are discussed here in order to bring forth the significance of this research to the field of assessment of emergent bilingual students.

### **Accountability: The Context for Assessment**

Currently, assessments have a growing role in national educational reform. They undergird both evolving teacher and student accountability plans. The promise of accountability to improve education is based on the commitment to “greater public responsibility for improving public education and greater equity of educational outcomes” (Guzenhauer & Hyde, 2007: 489). The inclusion and documentation of the performance of “sub-groups,” such as emergent bilinguals, in assessments is thought to drive reform by directing resources to the groups with greatest need. It is also assumed that inclusion of sub-groups of students in assessments will also allow for monitoring of their progress (Wang, Beckett & Brown, 2006).

The federal government has taken on an increasingly greater role in local educational practices primarily through legislation which ties funding to data-driven accountability. No Child Left Behind (NCLB) ushered in a wave of high-stakes testing and corresponding accountability measures more than a decade ago (2001).<sup>5</sup> Race to the Top (RTT) federal grants continue in this vein. In fact, the pending revised version of the NCLB legislation contains further assessment-based accountability measures that will impact emergent bilingual students (Lesaux, 2011).

---

<sup>5</sup> No Child Left Behind (2001) is most commonly known for two major policies. One is the testing requirements it sets forth for students and schools – yearly testing in English and Math, the inclusion of English Language Learners, the reporting of disaggregated data by sub-groups and meeting benchmark goals (Annual Yearly Progress) for student performance, among others. Secondly, Title II of NCLB, the Improving Teacher Quality program, stipulates that federal funding can also be used to support teacher learning which is based on acquiring, “the necessary subject-matter knowledge and teaching skills” in a given content area (Cochran-Smith & Lytle, 2006: 671).

Although data-driven accountability is coupled with the goal of educational equity, many researchers question whether test-based accountability, if not properly implemented, may contribute to greater disparities between sub-groups of students (Darling-Hammond & Falk, 1997). The breakdown of the logic of test-based accountability occurs because many assessments and the way they are interpreted lack the power to take into account diverse assessment needs of students, under which emergent bilingual students fall. Not taking into account these students' assessment needs may result in educational programming choices geared to raise test scores rather than to address the complex learning needs of a diverse body of students (Menken, 2008; Diamond & Spillane, 2004).

This research aims to further examine how assessments, both large-scale and classroom-based, impact the knowledge of elementary school bilingual teachers and the practices they implement. This research for this study was conducted in New York City. Although New York City will provide the context of the study, the issues of assessment for emergent bilingual students is similar regardless of geographical area because teachers are being asked to use assessments in similar ways across the nation under the framework of NCLB and RTT. Furthermore, the educational concerns regarding emergent bilingual students are not solely localized to states and urban areas traditionally thought of as having high numbers of these students such as California, Texas, New York and Florida. Rather, the emergent bilingual population has grown at high rates in states such as South Carolina and Kentucky, states not customarily thought of as having large numbers of these students (García, Kleifgen & Falchi, 2008). Although this doctoral research examines how teachers develop knowledge about emergent bilingual students and how they fashion instructional practices for them based on

assessments within the New York City educational context, I also hope the findings are relevant across contexts.

### **New York City’s Emergent Bilinguals: A Critical Piece in New York City Accountability Reform**

The New York City Department of Education (2010) defines emergent bilinguals, who they term English Language Learners as “students who come from homes where a language other than English is spoken and who score below a state-designated level of proficiency on a test of English language skills.” García, Kleifgen & Falchi (2008: 9) report that students considered emergent bilinguals are entering New York City’s public school system at a noteworthy pace and comprise a sizable portion of the entire student population. In fact, according to the Home Language Identification Survey (HLIDS)<sup>6</sup>, more than 41% of the entire city’s student population speaks a language other than English in their homes (NYCDOE, 2011). English Language Learners comprise 14.3% of the entire city’s public school population (NYCDOE, 2011). Therefore, bilingualism, more than being incidental, is quite close to the norm for students in the city’s public schools. In an effort to address the needs of emergent bilingual students, New York City schools host various programs for these students.

According to the New York City Department of Education, there are three main ways that emergent bilinguals are instructed in schools – English as a Second Language (ESL), Transitional Bilingual Education (TBE) and Dual Language Bilingual Education (DL). Each of these programs employs different approaches to instructing students and, as such, also has diverse goals.

---

<sup>6</sup> The Home Language Identification Survey (HLIDS) is filled out by all parents before registering their child into a New York City public school.

Of these three different programs available to emergent bilingual students in New York City, the majority of emergent bilinguals receive English language support through ESL – 70.2% (NYCDOE, 2011). ESL is a monolingual model of education whereby a teacher works with a group of students, either by “pulling out” students to a separate classroom or by “pushing into” a classroom – to provide English language instruction. Although students’ home language may be used within ESL, its use is rare. ESL instruction employs English as the main vehicle for instruction. The goal of this program is to support and strengthen students’ English literacy (García, 2009).

About twenty percent of the city’s emergent bilingual population is in transitional bilingual education (TBE) programs (NYCDOE, 2011). According to data published by New York City Department of Education (NYCDOE) during Spring 2010 there were 331 TBE programs across New York City public schools. Approximately 81% of these were Spanish TBE programs. The goal of TBE programs is to “transition” students from their home language to English. This model differs from ESL in that students are grouped with other students who speak the same home language into classes where this language is used to instruct students for a period of time until they are deemed English proficient according the state proficiency exam given yearly.

Lastly, 3.8% of emergent bilingual students are in dual language bilingual programs. Dual language refers to bilingual programs in which the student composition is equally divided between students who are learning English and those who are proficient in English. The motivation behind this configuration is that through interaction, students who use their home language will be language models and support for those students learning their second language.

However, the issue of program definition is a complex one. In many dual language bilingual programs, the student composition does not match the intended configuration. Rather, in this study, the majority of students enter into dual language bilingual programs have Spanish as a home language. These programs would better be described as developmental bilingual programs (García, 2009). Developmental bilingual programs are distinguished by having a majority of students that have the same home language. Students are taught literacy initially primarily in their home language with the end goal of biliteracy and bilingualism by the time of exit (García, 2009). In general the goal of dual language bilingual programs is to promote biliteracy and biculturalism for students across the continuum of bilingualism. The most current report issued by the NYCDOE states that there are currently 92 operating dual language bilingual programs across the city (NYCDOE, 2010). Although the majority of these are Spanish-English dual language bilingual programs (83.7%), there are also programs in Chinese, Haitian Creole, Russian, French and Korean (NYCDOE, 2010).

In the qualitative portion of this study, solely dual language bilingual teachers were interviewed. There are two primary reasons for this sample being entirely comprised of dual language bilingual teachers. The first one being during the recruitment process many principals that I contacted revealed that they did not have transitional bilingual classes in grades 3-5, thus leaving only dual language bilingual schools as sites where to conduct interviews. Secondly, dual language bilingual school principals may be open to having research conducted in their schools and with their teachers because along with research there may be the possibility of improving the assessment processes at their schools.

For the quantitative portion, teachers who work within all three types of programs for emergent bilinguals were surveyed. Despite the differences in programs and approaches to teaching emergent bilingual students, grade 3-5 teachers who work within all of these program models face the same expectations for using assessment to guide their instructional decisions. Therefore, it is important to understand the overarching framework of accountability in New York City and how it affects the assessment of these students.

While New York City's accountability system has its own particular qualities, the larger assessment structure set by New York State is ultimately determined by the federal mandate to administer standardized tests yearly from grades 3-12. Under NCLB, in New York State, like all other states, emergent bilingual students must take state exams after being in the country for a year or more. In New York State and elsewhere, the inclusion of emergent bilingual students in accountability systems has been hailed as a positive step towards ensuring that these students' performance and progress is held as a priority (Abedi, 2004). It is believed that the inclusion of emergent bilingual students into these assessment measures will provide information about these students' skills and abilities, hold them to high standards, and provide a catalyst for education improvement to close the achievement gap. Although emergent bilingual students have been added to the accountability mix, it is not clear how educators can be engaged to use data effectively to support instruction that is specifically targeted to meet the needs of emergent bilinguals.

In addition to New York State exams, New York City has developed its own accountability system comprised of both large-scale standardized tests and classroom-based assessments which gauge the performance and progress of all students. Each school is evaluated

in part by assessing how sub-groups of students, under which emergent bilingual students fall, perform and progress along standardized measures for English Language Arts. Furthermore, New York City public schools are required to define formative (on-going) assessments that will be used by teachers throughout the year to assess the progress of all students. Schools are evaluated periodically with respect to student performance, progress and use of data from both these types of assessments – large-scale standardized tests and classroom-based assessments. Given this framework of assessments coupled with the sizeable percentage of emergent bilinguals in the total student population, the performance of emergent bilinguals on these assessments is critical to a given school’s success under New York City’s accountability system.

New York City’s public schools provide a good example of how accountability, which includes both summative and formative assessment, impacts both students’ learning and teachers’ work. The example provided by New York City is indicative of a trend of the increasing intensity of testing, which is also seen in other cities and states.

The emphasis on testing has transformed students’ experience of school; tests have become a central part of schooling. For example, in New York City fourth graders have at least 14 days of mandatory testing (2 days of English Language Arts testing , 2 days of Mathematics testing, 2 days of ‘predictives’ in English Language Arts, 2 days of ‘predictives’ in Mathematics, 2 days of Science testing).<sup>7</sup> This amounts to 5.5% of all school days dedicated to testing. Furthermore, this does not include additional mandated testing for emergent bilingual students

---

<sup>7</sup> Predictive assessments are developed by Acuity for the NYCDOE in English reading and Mathematics. These assessments are given to students Grades 3-5. The content of these exams parallels the state exams that students in these grades will take. They are given twice a year before the official state exam is administered in order to provide timely information predicting how a given student will perform in the subsequent state exam. Teachers are provided with test results in order to develop ‘interventions’ for students.

such as the New York State English as a Second Language Achievement test (NYSESLAT), which would bring the percentage of the school days dedicated to testing for these students, closer to 8%.<sup>8</sup>

In addition, New York City schools are also mandated to implement and administer a school-based program of periodic assessments. These are meant to complement standardized data the school collects. Periodic assessments are formative in nature, meaning they are designed to be used during the year to guide instruction. Schools can choose between using a menu of periodic assessments provided by New York City Department of Education (NYCDOE) or develop their own. The amount of time allocated to periodic assessment varies according to the assessments chosen. Students may be assessed through assessments in test formats such as base-lines, mid-lines, and end-lines, unit tests and test simulations or through reading records, conferences and observations.<sup>9</sup>

---

<sup>8</sup> This tally was compiled from my own experiences as a teacher and a coach. This tally is only reflective of the 4<sup>th</sup> grade testing demands in New York City. Fourth grade students are the only ones in elementary school who take a state-wide science exam. For other grades these percentages may differ. However, they remain comparable. Starting in May 2011 in New York State, literacy and math exams are administered over 2 days for each content area and for all grades 3-5.

<sup>9</sup> Reading records are assessments that are administered to determine a student's reading level and comprehension. Conferences can take place during any content area. Teachers may have a variety of objectives when they conference with a student. If conferencing is tied to a school's periodic assessment, it is most likely tied to grade level content standards.

Base-lines, mid-lines, and end-lines refer to assessments which are meant to assess a range of student knowledge and content over a range of time. For example, a base-line is established on grade level standards of the previous year. Students are assessed along these standards to determine if they have acquired or if their knowledge exceed the standards to which students were held to at the end of the previous grade.

As exemplified above, both students and teachers are involved in assessment for a significant amount of time. These figures become alarming given that standardized assessments may not “fit” emergent bilinguals, and yet they are included in great numbers. Secondly, as this dissertation aims to investigate, it is not presently known what teachers do with either summative or formative data of their students. What is known suggests that the intensity and high-stakes nature of testing which has taken over the instructional agenda in many educational environments does not have positive effects on the emergent bilingual population (García, 2010).

These points taken together illustrate the high level of resources, both financial and in terms of time, which are dedicated to the administration of assessment of all students, including emergent bilinguals. It also exemplifies how few resources are directed towards examining how assessment can be translated to useable knowledge that impacts student learning, in particular for emergent bilingual students. In order for assessment to have any effect on emergent bilingual student learning, understanding how teachers successfully navigate and generate knowledge from various types of assessments is necessary. Therefore, this research examines what teachers learn about emergent bilingual students from these assessments and what instructional changes result from this knowledge. This is a good starting point to evaluate if current accountability measures are meeting their intended vision for emergent bilingual students.

## **CHAPTER 2:**

### **Literature Review**

#### **Introduction**

How teachers generate knowledge based on reading assessments and use this knowledge to develop teaching practices for emergent bilingual students is the focus of this study. This literature review examines how researchers have defined and studied the four key topics of this study: teacher knowledge, accountability, assessment of bilingual students, and bilingual reading development. The first section on teacher knowledge provides an overview of how teacher knowledge has been conceptualized and evaluates how current models of teacher knowledge accommodate the knowledge of emergent bilinguals. The following section on accountability defines formative and summative assessments, as well as provides a description about the current accountability context and its impact on emergent bilinguals. The third section offers a synopsis of research on how summative and formative assessments have been adapted to the needs of emergent bilinguals. The last section provides a brief overview about reading development in general and describes how emergent bilingual students develop reading skills differently than monolingual students.

#### **Part 1: The Evolution of Teacher Knowledge**

Teacher knowledge and practice are crucial in developing sound learning environments for students (Black, Harris & Lee, 2003, 2003; Falk, 2003; Heritage et al., 2009). Therefore, it is important to define what counts as teacher knowledge, how it is acquired, and what types of teacher knowledge are critical for the teaching of emergent bilingual students. Despite the widespread acceptance among scholars and educators that teacher knowledge is critical to solid

instructional practices, how it is conceived of is disputed (Cochran-Smith & Lytle, 2006; Shulman, 1986). In order to embark on the study of the development of teacher knowledge from assessment of emergent bilingual students, it is necessary to analyze existing models of teacher knowledge to determine the extent to which each of these models adequately capture how teachers may acquire knowledge from the assessment of emergent bilingual students.

Teacher knowledge is contested among a number of different fronts (Cochran-Smith & Lytle, 2006). *What* type of knowledge is recognized as most central to teaching and learning is one line along which there is dispute. One position is that content knowledge, or knowledge of a particular subject, is most relevant to effective teaching and learning. In fact, this viewpoint has a long-standing historical tradition. In his analysis of teacher examinations from 19<sup>th</sup> century United States, Shulman (1968: 5) concluded, “although knowledge of the theories and methods of teaching is important, it plays a decidedly secondary role in the qualifications of a teacher.” Content knowledge has played a central role in teacher education and has often been privileged above other forms of knowledge that teachers require.

Cochran-Smith & Lytle (2006) assert that the emphasis on content knowledge has indeed carried over until today and, in fact, has been renewed and strengthened by NCLB legislation. In their study of NCLB legislation’s stipulations on teacher education, they report that NCLB favors “subject matter knowledge” above others (Cochran-Smith & Lytle: 671). Their analysis of the implications of NCLB on teacher knowledge and qualifications points to a larger idea that, “knowledge is an object to be transmitted from outside experts to teachers, who then transmit it to students” (Cochran-Smith & Lytle: 674). This stance is aligned with the idea that teaching and student learning operate along one dimension – they are context and individual free.

Although no educator would dispute the importance of content knowledge, researchers have refuted the premise that content knowledge can stand alone.

What types of teacher knowledge have been put forth by researchers and how has the conceptualization of teacher knowledge evolved? One of the first models which categorized types of teacher knowledge was offered by Shulman (1986). This model posed a “new” branch of knowledge in addition to content knowledge. While content knowledge was thought to be the main, and perhaps, lone component of teachers’ store of information needed to teach, Shulman added a new domain entitled “teacher knowledge.” Therefore, Shulman’s (1986) model of teacher knowledge includes two broad domains of knowledge which describe the knowledge required to know *what* teachers will teach (content knowledge) and the knowledge required to inform *how* to teach it (teacher knowledge). Both of these domains – content knowledge and teacher knowledge – can be divided into useful subcategories (see Table 1). While the Shulman’s label of “teacher knowledge” is initially vague, his subcategories bring more clarity to this category. They demonstrate the importance of teacher knowledge that emerges from experience.

**Table 1. Shulman’s Model of Teacher Knowledge (1986)**

<b>Content Knowledge</b>	<b>Definition</b>
<i>Subject-Matter Knowledge</i>	“Amount and organization of knowledge per se in the mind of the teacher” (p. 9).
<i>Pedagogical Content Knowledge</i>	Understanding of the “teach-ability” of the subject-matter.
<i>Curricular Knowledge</i>	Understanding of the program of study.
<b>Teacher Knowledge</b>	
<i>Propositional Knowledge</i>	“Accumulated wisdom of practice and in many cases is as important a source of practice as theory or empirical principles” (p. 11).
<i>Case Knowledge</i>	“Knowledge of specific, well-documented, and richly described events” (p. 11).
<i>Strategic Knowledge</i>	“Comes into play as the teacher confronts particular situations or problems, whether theoretical, practical or moral, where principles contradict one another, or the precedents of particular cases are incompatible” (p. 13).

Both of Shulman's proposed types of knowledge are academic and practical. Although teachers' context-bound experience is included in this model, how teachers develop knowledge of students is not explicitly addressed. For example, Shulman's category of teacher knowledge describes how a given teacher's knowledge of content shifts through the act of teaching, but does not indicate the central role that understanding students – their developmental needs, their cultural and linguistic background – has in enriching a given teacher's knowledge base. Nonetheless, Shulman's model gives importance to teacher knowledge that is learned and applied *through* the craft of teaching, thus parting with previous views that content knowledge, divorced from practice, is the primary factor in a teacher's success. The addition of this element, teacher knowledge that is developed through context, is a foundation for further models which build on the idea that knowledge emerges from teachers' experiences.

Another model, the Framework for Teaching by Danielson (1996), provides a detailed breakdown of the different types of teacher knowledge. Danielson's framework was designed for practical purposes, specifically teacher self-reflection. This model includes a wider range of components for teacher knowledge than does Shulman's framework, including the addition of teacher knowledge of students (see Table 2).

**Table 2. Danielson’s Framework for Teaching: Focus on Teacher Knowledge (1996)**

<b>Demonstrating Knowledge of Content Pedagogy</b>	<b>Definition</b>
<i>Knowledge of Content</i>	“Sufficient command of a subject to guide student learning.” “Content encompasses all aspects of a subject: concepts, principles, relationships, methods of inquiry and outstanding issues” (p. 62).
<i>Knowledge of Prerequisite Relationships</i>	Understanding of what students need to know for student learning to occur.
<i>Knowledge of Content-related Pedagogy</i>	Understanding of how to enact pedagogy that best suits the learning of a given content area.
<b>Demonstrating Knowledge of Students</b>	
<i>Knowledge of Characteristics of Age Group</i>	Understanding of developmental age characteristics of student group.
<i>Knowledge of Students’ Varied Approaches to Learning</i>	Understanding of the varied learning styles that students have.
<i>Knowledge of Students’ Skills &amp; Knowledge</i>	Understanding of students’ level of skills and knowledge and how this information impacts instruction.
<i>Knowledge of Students’ Interests and Cultural Heritage</i>	Understanding of students’ interests and cultural heritage and how this information impacts instruction.

Shulman’s and Danielson’s models are similar in recognizing the importance of content and teacher knowledge. Danielson’s model clearly evolves from the initial categories developed by Shulman. She combines Shulman’s content and teacher knowledge into one category entitled, “knowledge of content pedagogy.” However, Danielson provides a novel category of “demonstrating knowledge of students” which includes knowledge about child development, culture and individual characteristics such as learning style, interests, and skills. The addition of “knowledge of students” reflects the shift from viewing teacher knowledge as a set of static skills to conceptualizing teaching as embedded within a changing context, where students are at the center.

Moreover, the purpose of this framework is teacher goal setting and growth. This purpose indicates a shift in the orientation of the models from educating teachers by providing

them with requisite knowledge to viewing teachers as essential in directing their own development. The character of this model is complementary to the goals of the proposed research which aim to describe the development of teacher knowledge as it emerges from practice. Danielson’s model represents a change in thinking about teacher knowledge – that it is acquired through experience and reflection. This model demonstrates an evolution in redefining teachers in relation to their knowledge, making them agents in directing their own learning.

With respect to the proposed research, a model that addresses what knowledge teachers need in order to work with assessments is critical. Heritage (2007) proposes such a model: it offers four different types of knowledge that are required by teachers in order to put information from assessments to use (see Table 3).

**Table 3. Heritage’s Types of Teacher Knowledge Required for Assessment (2007)**

<i>Domain Knowledge</i>	Understanding the progression of concepts or skills within a particular topic (domain).
<i>Pedagogical Content Knowledge</i>	Knowledge required for teachers to identify multiple ways to teach students within a specific topic or domain.
<i>Knowledge of Students’ Previous Learning</i>	The ability to tease out students’ “level of knowledge in a content area, understanding of concepts in the content area, the level of their skills specific to the content area, the attitudes the students are developing and their level of language proficiency” (p. 142).
<i>Assessment Knowledge</i>	How assessments can lead to practices which support students (p. 142).

This model builds from the previous two models in keeping content knowledge (here referred to as domain knowledge), pedagogical content knowledge, and knowledge of students. Heritage’s model presents two innovations – the addition of ‘language proficiency’ within the category of “knowledge of students’ previous learning” and of “assessment knowledge.” The addition of these components reflects the changing landscape of teaching. As the student population changes and as policy changes educational priorities, teacher knowledge must also

reflect these shifts. Assessment knowledge refers to not only understanding how to effectively administer assessments, but also what practices can emerge from assessment to further student learning. Additionally, although knowledge of students has been conceptualized in previous models (Danielson, 1996; Shulman, 1986), the inclusion of language proficiency is an important one, as it points to the recognition of how language proficiency may influence student performance. Heritage's (2007) model reflects the evolution from Shulman's (1986) and Danielson's (1996) models through increasingly incorporating contextual factors, such as the teaching environment, student characteristics, student diversity, and ultimately policies into the models of teacher knowledge.

With respect to the model of teacher knowledge that will be used for this study, I will start with an examination of Heritage's since it is the most current and incorporates assessment. Heritage's (2007) model of teacher knowledge was created in relation to formative assessments. As to how this model can be applied to the dissertation research, the question of whether the four strands of knowledge (domain, pedagogical content, students' previous learning, and assessment) are required for the analysis and use of data, summative assessments must be considered. Since summative assessments are different in nature from formative assessments in their development and purpose, the ways in which teachers learn from these types of assessments may differ substantially from how they learn from formative ones. For example, with respect to assessment knowledge, teachers may need to understand what a particular formative assessment reveals about how students process a specific concept, while the knowledge necessary for a teacher to make use of the results of a summative assessment may include how to interpret scaled scores and proficiency levels.

Finally in Heritage's (2007) model, although student language proficiency was included in knowledge of students' previous learning, this component alone may not be sufficient to describe the knowledge base that teachers of emergent bilingual students need in order to adequately instruct them. I propose that teachers of bilingual students also require knowledge of bilingualism and language development in order to transform knowledge gleaned from assessments into well-developed practices for these students. In proposing this type of knowledge, I offer a brief review of some basic concepts of bilingualism in order to contextualize why knowledge of students' language history, language proficiency and bilingualism are important with respect to assessment.

Bilingualism is more than knowing two languages in isolation from each other; bilingualism exists on a continuum and is dynamic (García, 2009; Valdés, 1994). García (2009: 53) writes, "bilingualism is not simply linear but *dynamic*, drawing from the different contexts in which it develops and functions." Furthermore, Valdés (1994) suggests that bilinguals can function in *more than one* language. Both of these ideas have significant implications for the assessment of bilingual students in general. Bilingual students can learn and comprehend in two languages. However, how accurately bilingual students express what they know, may depend on the language in which they operate. Teachers who work with bilingual students in general, and with emergent bilinguals, in particular, must be aware of how students' language history and proficiency may influence their experience of that assessment and subsequent results.

Therefore, for this dissertation research, Heritage's (2007) model of teacher knowledge will be adapted with bilingualism in mind. I propose a model that includes the addition of *bilingual knowledge* and replaces domain knowledge (see Table 4). Although, this additional

category of bilingual knowledge is complementary to the category of “knowledge of students’ level of language proficiency” found in Heritage’s original model, these two categories are different in that, bilingual knowledge refers to the information teachers need in order to understand the varied ways in which students develop and learn in two languages. In contrast, knowledge of students’ level of proficiency solely looks at “their ability to speak, listen, read and write in the L2 [second language] for both conversational and academic purposes” (Celic, 2009: 36). Thus, the additional category of bilingual knowledge is critical to describing what and how teachers conceive of students’ bilingual literacy development. In this model, domain knowledge is excluded because this study is focused on teachers’ applied knowledge within classrooms and does not aim to measure their general knowledge of reading. Furthermore, the instruments designed for this study, as will be described in the methodology section, are not adequate the extent of teachers’ knowledge of reading practices.

**Table 4. Ascenzi-Moreno’s Model of Knowledge for Bilingual Teachers**

<i>Bilingual Knowledge</i>	Knowledge of how two languages develop simultaneously and how to facilitate the development of language and content in two languages.
<i>Student Knowledge</i>	The ability to tease out students’ “level of knowledge in a content area, understanding the concepts in the content area, the level of their skills specific to the content area, the attitudes the students are developing and their level of language proficiency” (Heritage, 2007: 142)
<i>Assessment Knowledge</i>	How assessments can lead to practices which support students.
<i>Pedagogical Content Knowledge</i>	Knowledge required for teachers to identify multiple ways to teach students within a specific topic or domain.

The identification of components of teacher knowledge is a necessary precursor to studying how different aspects of teacher knowledge may shape instruction. Although empirical studies of how teacher knowledge impacts student learning is in its infancy, the available

research does point to the importance of teacher knowledge in impacting student learning. For example, McCutchen, Green, Abbott & Sanders (2009) examined how elementary school teachers' level of knowledge impacted students' reading outcomes. They found that teacher knowledge, particularly linguistic knowledge, had a positive impact on student outcomes for all student populations examined in the research, but had the greatest effect on low-performing students. Yet, teacher content knowledge is only one component of the necessary body of knowledge required to effectively use assessments (Heritage & Chen, 2005; Heritage, 2007; Heritage et al., 2009). This limitation has been noted by researchers who have called for more research on how to operationalize other forms of teacher knowledge – such as pedagogical knowledge – and use these to further explore how multiple facets of teacher knowledge impact instruction (Carlisle, Correnti, Phelps & Zeng, 2009; McCutchen et al., 2009). In this research I attempt to understand what types of knowledge teachers of emergent bilinguals report learning from assessments of emergent bilinguals and how these different forms of knowledge impact instruction.

In the following section, I offer definitions of formative and summative assessments in order to provide an overview of the assessments emergent bilingual students are administered. I also provide a description of how accountability measures currently direct school change.

## **Part 2: Assessment & Accountability**

### **Assessment**

The word assessment is derived from the Latin word, *assedere*, which means “to sit with” (García, 2009). The origin of this word evokes a relationship – the teacher “studies” the student to understand both what the student knows and how the student comes to know. This

relationship is centered on the student's emerging knowledge base and the teacher's emerging knowledge of the student. The word assessment today encompasses both the vision described above, as well as a broad range of activities that include individual teacher conferencing to large-scale standardized assessments (García, 2009). In this review, the terms "formative" and "summative" will be used to describe the function that assessment plays with regards to teacher knowledge and instruction. These terms will be described with more detail in the following paragraphs.

Assessment can be used for multiple purposes, at multiple levels of educational decision-making and by a variety of stakeholders. Ideally all assessments are administered for the purpose of making instructionally meaningful decisions, whether they are to be used to steer classroom instruction or in the creation of policies for groups of learners at the system-wide level.

For example, a classroom observation of an individual student can be used to develop instructional priorities for that student and may inform the teacher's understanding of others at the classroom level. However, an observation of an individual student will provide limited information as to how to direct educational decisions for an entire school population. Correspondingly, results from a large-scale standardized test may provide useful information about the learning trends of large groups of students, but may not yield valid inferences about an individual student. Teachers currently work with these two types of assessments in the expectation that analysis of the data will drive their instructional decision-making. This dissertation research asks if teachers learn from either type of assessment, thus providing a

holistic perspective on the impact of these measures on teacher knowledge of dual language bilingual teachers.

One way to distinguish between the two types of assessments described in the previous paragraph is to classify them as either formative or summative (Black et al., 2003; Gottlieb, 2006; Heritage & Niemi, 2006). Heritage & Niemi (2006: 265) write, “formative assessment is a systematic process to continuously gather evidence and provide feedback about learning while instruction is underway.” Teachers use formative assessment to adapt or redirect instructional moves based on the assessment (Black, Harris & Lee, 2003). Formative assessment is at the heart of informed decision-making and is based on student’s demonstrated learning. Popham makes the case that formative assessments are the first step in a learning progression (2008). He defines learning progressions as the “step-by-step” building blocks students need in order to successfully attain a more distant, designated instructional outcome (Popham, 2008: 24). In other words, formative assessments are the starting point for fashioning instruction for students by providing insight into the boundaries of students’ knowledge and skills.

Various types of formative assessments include running records of student reading, use of rubrics, and checklists to assess student work and observations of students, among others. Formative assessment is classroom-based — it is administered to students by their classroom teacher or another teacher that works with them and used by these teachers. In this research, the focus will be on running records, an assessment method that teachers use to evaluate individual student reading.

In contrast, summative assessments provide an indicator of student learning through reporting of student achievement at the end of unit of study or school year (Gottlieb, 2006).

These assessments may take a variety of forms, such as exit interviews, end of unit tests, final exams, benchmark tests, and standardized tests. Assessments are considered summative if they *sum up* what a student has learned within a certain period of time. Summative assessment is generally used to grade individual students or provide information for evaluating classes, schools or school systems. Since they summarize large amounts of information, Koretz (2009) maintains that large-scale standardized tests are simply assessments that reflect a sample of the knowledge and skills of a given content area. Instructionally, standardized tests are used as benchmarks in order to provide school administrators and teachers with an overview of student performance at the end or beginning of the school year. In this thesis, the knowledge gained from large-scale standardized literacy test data will be contrasted with that from running records, a classroom-based assessment often used by teachers.

In sum, formative assessments are designed to provide on-going information about student abilities so that teachers can make decisions about how to direct instruction. Summative assessments are designed to provide a pulse of student performance by assessing students on a representative sample of questions. The results are used to assess the acquisition of a given learning goal and move a student forward for formative assessments or, as in the case of large-scale standardized tests, to provide a comparative snapshot of trends of performance across and within populations of students (Popham, 2008).

As stated in the introduction, teachers in this study frequently work with the data of emergent bilingual students from both of these types of assessments. The learning that occurs as a result of teachers' work with sets of data of emergent bilingual students from either assessment is not entirely clear. In fact, the literature on teacher knowledge generated from assessment in

general is slim (Heritage et al., 2009). Alternatively, the research on test-based accountability measures on student achievement and school change has been more robustly studied. I now present a definition of accountability and examine the impact accountability has had on instruction.

## **Accountability**

Although testing and accountability are often equated with each other, they are different. Accountability stands for the “implicit or explicit action theories that determine to whom, for what and how” schools must demonstrate their responsibility of successfully educating students (Elmore, 2005). Accountability systems are designed to improve the instruction that students receive and enhance student learning. Elmore asserts that school accountability has been ever present and has taken different forms. Elmore proposes that accountability is experienced by school leadership in three basic ways (2005). The first filter through which accountability is experienced is *individual responsibility*, or school leaders’ beliefs about education, student learning, the role of teachers, and their role in supporting the educational process. Second, *collective expectations* are defined as the practices that guide what teachers and students do. Lastly, Elmore proposes *alignment* as the process of ensuring that individual responsibility and collective expectation are coherent and consistent. In Elmore’s model of accountability, school leaders, major decision-makers in how assessments are to be used by teachers, can be accountable to both individual responsibility which includes the needs of teachers and students, as well as collective responsibilities which may include testing mandates. In fact, his category of alignment suggests that there is a balance between the two, rather than an exclusive focus on one or the other. Other researchers have also provided

explanations of accountability in order to provide an alternative perspective to the current use of the term.

Linn asserts that accountability must be shared among students, teachers, administrators and policymakers (2003). As such, “it is the most likely way for expectations [for school improvement] to be achieved” (Linn, 2003: 3). Linn also questions the narrowness of current accountability measures which rest almost exclusively on high-stakes tests. He suggests that the unbalanced focus on high-stakes testing renders a limited vision of what factors contribute to student improvement. Linn suggests that shared accountability systems that incorporate a broad definition of what counts have the potential to “enhance positive effects and minimize negative effects” of accountability systems (2003: 3). However, these definitions and descriptions of accountability by Elmore (2005) and Linn (2003) are far from the ones currently in place.

No Child Left Behind (2001) has provided the backdrop for the majority of the testing now occurring in schools. In the current climate of accountability, standardized tests have taken center stage. In fact, accountability has been simplified to be based solely on data from standardized assessments. As Linn (2003) contends, the danger of an exclusive focus on testing is that teachers narrow their instructional priorities. This negative effect stands in odds with the original intent of this legislation, which is to close the achievement gap between groups of students. It was thought that including sub-groups of students, such as emergent bilinguals into the testing pool, was one way to close the achievement gap in that the inclusion of these “subgroups” of students in testing would be a step in ensuring equity. As Abedi, Hofstetter & Lord (2004: 3) note, “the exclusion of certain groups from large-scale assessments may have

been well-intentioned with regard to fairness and validity, but it has resulted in a lack of representation in broader educational policy and accountability debates and thus has diminished educational opportunities for English learners.” This charge of improving educational opportunity through test-based accountability is a noble one. However, research findings show that this goal of test-based accountability measures has thus far fallen short.

Menken’s (2008) study of testing and language policy in New York City high schools demonstrated that the priority given to emergent bilinguals’ performance on tests has deeply impacted the teaching and learning opportunities that these students received. The study illustrated in detail how teachers of emergent bilingual students focused instruction on skills that students would need for performing well on the exam, rather than on practices that would benefit these students’ learning in the long run. Emphasizing the skills that students needed for the exam set bilingual students’ actual instructional needs at a lower priority. Her study clearly showed that the effects of standardized testing on bilingual students stands in contrast to the intended effect of NCLB which aims to raise achievement of all students by including them in annual testing and monitoring their progress (Ladd, 2008).

In fact, Ladd (2008) states that accountability measures through high-stakes testing have not facilitated substantial gains in student achievement overall. Rather, there is a risk that accountability systems which are almost exclusively focused on testing may actually encumber educational improvement (Linn, 2003). Linn (2003: 8) notes the effects of accountability on a system-wide level; he states, “for schools that continue to fall into the improvement category may actually hinder educational excellence because they implicitly encourage states to water

down their content and performance standards in order to reduce the risk of sanctions for their schools.”

This school-level impact is well documented by other researchers. Diamond and Spillane (2004) find in their comparative study of the effects of accountability measures between the lowest performing schools and higher performing ones that schools respond differently to high-stakes policies testing policy. Lowest performing schools directed their efforts “in ways that were designed to respond to the policy demands of the external environment – getting off probation” (1165). Therefore, these schools did not engage in school-wide educational reform to increase student achievement of all students, but rather focused energies on raising the performance of a small segment of low-performing students.

Teachers also are negatively affected by the demands of test-based accountability measures. In fact, although the amount of time a teacher dedicates to assessment is difficult to calculate, it is a central component of professional responsibility. Furthermore, this emphasis on assessment “from the outside” has clouded teacher’s understanding of their role in assessment. Heritage (2007: 140) writes, “teachers do not control how or when these tests occur, what the purpose of the assessment is, or who is assessed.” As such, teachers’ understanding of their role in using assessment, even formative ones, to strengthen teaching and learning has been lessened.

These examples highlight the limits of test-based accountability measures to raise educational achievement in general. These studies also raise the importance of examining how accountability systems affect educators’ knowledge base and their practice. Gunzenhauser & Hyde (2007) speak to the pivotal role of educators within an accountability system. These authors quote Elmore, “accountability systems themselves do not directly ‘cause’ schools to

increase the quality of student learning and academic performance. At best, they set in motion a complex chain of events that may ultimately result in improved learning and performance (cited in Gunzenhauser & Hyde: 503).”

The promise that assessment holds for educational equity resides in the ability of a school community and, ultimately, teachers to improve student learning through a variety of means. The identification of student achievement levels from assessments alone cannot direct improved student learning. Rather, teachers must be able to use their understanding of students’ knowledge to base their next instructional steps. Assessment and instruction are not the same, but rather are complementary. Bachman (2001: 298) suggests, “we cannot lull ourselves into believing that the concerns of measurement are exactly the same as those for instruction.” In other words, assessment is not an instructional strategy, but rather an educational practice that exists in a reciprocal relationship with instruction. Therefore, it is important to understand the relationship between assessment and improved student learning. In this vein, Heritage et al. (2009: 24) write that teachers need to, “infer the gap between the students’ current learning and the desired instructional goals, identifying students’ emerging understanding or skills so that they can build on these by modifying instruction to facilitate growth.” In addition, teachers must also use assessments to establish appropriate goals for students or cohorts of students, allocate support services for students, and effectively communicate student learning to parents (Darling-Hammond & Falk, 1997; Sato et al, 2008). Therefore, the accurate interpretation of assessments by teachers is vital to ensuring that student learning is appropriately supported through instruction, resource allocation, and goal setting, particularly when addressing the needs of a student population with specialized needs, such as emergent bilinguals. In this climate of

accountability, assessment is viewed as the reform. Yet, assessment alone is unable to affect instructional change. It is only through teachers that the promise of assessment-based accountability measures can be realized. I now turn to a discussion of the research that has been done on how to adapt or modify the assessments of emergent bilinguals.

### **Part 3: Assessment of Emergent Bilinguals**

#### **Bilingualism & Assessment of Content**

This section will provide an overview of the relationship between bilingualism and assessment of content. The term content is used here in order to denote the assessment of subject areas such as reading, social studies, and math. In the case of this research, the content being assessed is reading.<sup>10</sup> It is also used in order to make a distinction from the assessment of language proficiency. I begin by describing how the testing needs of emergent bilinguals differs from those of the monolingual English population and proceed to outline how assessment policy affects the education of bilingual students. It is important to note that many of the issues regarding the relationship of bilingualism and assessment have been explored from the perspective of summative assessment, in particular large-scale standardized tests. Therefore, there is limited information about how the practices developed to render standardized tests fairer for bilingual students apply to formative assessments.

Bilingual students are assessed for a variety of purposes. These purposes include determining their proficiency level in either of their languages, assessing their level of content

---

<sup>10</sup> Reading is not usually considered a content area. However, elementary school teachers teach reading as a subject as students are learning to put together a range of skills necessary to read (decoding, fluency, comprehension). Therefore, in this dissertation reading is considered a content area.

knowledge, or establishing whether they qualify for specific services or programs. Despite these diverse purposes, bilingual students often encounter a singular *method* of assessment – a monolingual one (García, 2009). Yet, the assessment of language proficiency is probably the only purpose for which a monolingual assessment is required. Assessment for other purposes, such as content, becomes problematic because bilingual students develop and retain information and knowledge in two languages (Valdés & Figueroa, 1994). Bedore, Peña, García & Cortez (2005: 190), note that, “single-language measures ignore the fact that bilingual children may choose to use different words depending on the setting, interlocutor, and context.” As a result, knowledge which emergent bilingual students possess and are able to demonstrate is bounded by their experience of the content in a specific language and/or by their language proficiency. Therefore, a bilingual student’s knowledge base may overlap between both languages or may solely be accessed in one language or the other depending on experience and proficiency.

Bilingualism is contextual and individual-specific (Valdés & Figueroa, 1994). Bilingual students are diverse. Their diversity spans race, ethnicity, educational history, economic status, and language history, among others. The unique mixture of these contextual factors impacts bilingual students’ level of proficiency and use of language. It is logical that students, who are at the beginning stages of proficiency in one language or have not acquired academic language through their educational experiences, may not be able to express their knowledge on an assessment, even if they possess it. These elements – race, ethnicity, educational history, economic status and language history – all make up a bilingual learner’s profile and may affect their performance on assessments.

Aside from proficiency, there may be other factors which affect bilingual students' performance. Butler & Stevens (2001: p. 420) state that "opportunity to learn" may be a strong influence on the performance of emergent bilingual students." Opportunity to learn in its simplest sense refers to sufficient time and adequate instruction to facilitate learning (Abedi, Courtney, Leon, Kao & Azzam, 2006). It has also been expanded to include the constructs of "content coverage, exposure, emphasis, and quality of instructional delivery," as well as level of cognitive demand of the instruction (Abedi et al., 2006). The concept of opportunity to learn is pertinent to the assessment of emergent bilinguals in assuring that they have had proper exposure to a given area being assessed. Assessment is based in the premise that students assessed have received instruction in the area evaluated (Butler & Stevens, 2002). Therefore, in addition to language proficiency, opportunity to learn may be a powerful contributor to an emergent bilingual student's performance on assessments. Opportunity to learn may influence how students in low-income areas perform in general (Butler & Stevens, 2002). However, in combination with emergent bilinguals' beginning language proficiency, these two factors together may affect student performance on assessment to an even greater degree than if language proficiency were the sole issue at hand. As a result, assessments of emergent bilingual students are not clear cut and may not truly reflect students' skill levels in a given content area. Not surprisingly, for emergent bilingual students, assessments are a less than perfect measure of their knowledge and skills (Abedi et al., 2006).

This raises the question: If the proper instruments to assess bilingual students are not available, why include these students in assessment campaigns? Emergent bilingual students were excluded from standardized testing until the 1994 Reauthorization of the Elementary and

Secondary Education Act (ESEA). This legislation stipulated that English Language Learners meet state-wide standards (Menken, 2008). One way of ensuring that students met these standards was through assessment. NCLB, another reauthorization of the ESEA, went further in mandating assessments of proficiency and content, as well as designating progress benchmarks for English Language Learners. These two pieces of legislation were devised to address the two main policy problems associated with the previous exclusion of emergent bilingual students from standardized testing. The first problem was the lack of information about how emergent bilingual students fared at a specific point in time, as well as longitudinally. Rivera (in Lachat, 1999: 1) referred to this as “systematic ignorance” about the performance and progress of bilingual students. Second, it was believed that the exclusion of emergent bilingual students from standardized assessments denied them from being held to rigorous standards. Coltrane (2002: 1) notes that the exclusion of bilingual students barred them from educational initiatives intended to raise achievement which were implemented for other student populations; she writes, “they [emergent bilinguals] have not reaped the benefits of educational initiatives and reforms intended to raise academic standards and promote student learning.”

Critiques about the exclusion of emergent bilinguals from wide-scale assessment that emerged in the late 90s are now moot – emergent bilinguals are regularly included in assessments in all states as a result of NCLB’s testing requirements. The current and most pressing issues regarding the assessment of emergent bilinguals are how to most appropriately assess them and what to do with the data that results from these assessments. The issue of appropriate assessments is a necessary precursor to any viable assessment policy. Resnick (2006: 36) notes, “they [accountability policies] cannot have their intended effect on the quality

and accessibility of education unless other components of the standards system – especially tests and measurements – are brought up to par,” thus underscoring the pressing need to develop assessments that are appropriate for emergent bilingual students to use.

Assessment policy – what gets assessed, by what means, and how it is assessed – affects the practice of teaching. However, assessment practices have reflected the belief that emergent bilingual students will test the same way as other students. According to La Celle-Peterson & Rivera (1994: 56), “the implicit guiding assumption appears to be that whatever curricular revisions and/or assessment innovations contribute to the success of monolingual students will also work for English language learners – that once English Language Learners know a little English, the new and improved assessments will fit them too.” Nevertheless, the instructional needs of emergent bilinguals are different from monolingual students. Lachat (1999: 51) notes, “assessment policies that were designed without the diversity of today’s population of English language learners in mind exert a powerful influence over every aspect of these students’ education”

Striking a balance between including emergent bilinguals in wide-scale assessment campaigns, while ensuring that the assessments which these students take accurately represent what they know is the focus of many current studies. The main thrust of current research on the assessment of emergent bilinguals is to examine how assessments can be made fairer for emergent bilingual students. This doctoral research acknowledges the difficulties in developing proper assessments for emergent bilingual students, however, the focus of this research is on examining how data from these assessments contributes to teacher knowledge and the potential impact of this knowledge acquired from the analysis of assessments of emergent bilingual

students on instructional practice. The following section of this review examines in what ways summative assessments fare with respect to evaluating emergent bilingual students' knowledge and skills.

### **Summative Assessments and Bilingualism**

For students whose first language is not English, assessments measure content knowledge and skills as well as *language*. The Standards for Educational and Psychological Testing developed by American Educational Research Association (AERA), American Psychological Association (APA) and National Council on Measurement (NCME) make the point that language is confounded with ability and skill for all tests that are language-based. The document states, “for all test takers, any test that employs language is, in part, a measure of their language skills. This is of particular concern for test takers whose first language is not the language of the test” (Abedi, 2008).

Yet, under the inclusion requirement under NCLB mandates, the testing of ninety-five percent of all students in a school is required, thus broadening the scope of who is tested to include bilingual students classified as English Language Learners. The performance and progress (difference in scores from one test administration to the next) of emergent bilinguals is critical to the strength of the accountability system, as this percentage holds considerable weight in determining if a school is deemed to be performing at appropriate levels. These two factors, the inclusion of emergent bilinguals into accountability systems and the high-stakes nature of student performance and progress, makes it of the utmost priority to question whether these standardized tests are reliable, and yield valid and unbiased information for the emergent bilingual students who take them.

Before reviewing the germane research examining the appropriateness of standardized tests for English Language Learners, brief definitions of reliability, validity and bias will be offered. Although these concepts will not be explicitly studied through this research, it is necessary to examine them in order to understand the critiques of large-scale assessment of emergent bilinguals.

Reliability refers to the ability of a test item or a test as a whole to consistently measure a designated construct or set of constructs (Koretz, 2008; Tobias, 2009). Validity is a sister concept to reliability. It refers to the ability assess what a given assessment is supposed to test. Chatterji (2003: 56) summarizes it as, “the *meaningfulness* of an assessment’s results for particular constructs, given the purposes for which the assessment is used and the populations relevant to the assessment context .” Therefore, validity requires that the test items have high reliability, but also that test results are analyzed appropriately.

Another important concept is bias. It is defined as, “a systematic distortion in scores that undermines the validity of a particular inference (Koretz, 2008: 262).” Koretz warns about concluding that differences between the performance of different groups is the result of bias. Rather, one way to ascertain if differences in performance between groups is the result of bias is by analyzing how students from different groups but of similar proficiency levels perform on a specific test item. This type of analysis is referred to as differentiation item functioning or DIF (Koretz, 2008). Determining that DIF exists does not explain why these differences between groups occur. Further analyses are required to determine the source of the DIF for students and ultimately if test items are biased for a particular group.

Applying the concepts of reliability, validity and bias to the current accountability context, brings forth the overwhelming challenge of choosing and administering standardized tests that are reliable, valid, and unbiased. This point calls into question the power of these assessments to influence instruction for students for whom the test may not be reliable nor valid. If we are unable to accurately infer what students are able to do from test results, then the results may not be valid for making sound instructional decisions. The majority of research on the assessment of emergent bilingual students has addressed the extent to which standardized tests are reliable and valid for this population. However, research has not focused on whether the actual or perceived level of validity and reliability of these exams for emergent bilingual students impacts what teachers learn about students and how they use this information to change instructional practice. Furthermore, previous research has not examined how teachers interpret or adapt the data from large-scaled exams to the specific needs of emergent bilinguals, if at all. The proposed study aims to contribute to understanding how teachers navigate through the data from this type of assessment for the benefit of emergent bilingual students.

Applying these definitions to the construction and use of summative assessments, such as large-scale standardized exams, reveals a number of challenges in capturing emergent bilingual students' knowledge. One of most salient problems is with the standardized character of many summative assessments. While students' language fall along a dynamic spectrum of language ability in English, test items are linguistically static. Therefore, students whose proficiency level is not commensurate with a *test-item's* linguistic demands may not demonstrate their knowledge of this item despite having the requisite content knowledge. Researchers have demonstrated empirically that English Language Learners' and non-English Language Learners'

performance differs in part based on the linguistic demands of a particular test item (Abedi, 2009; Martiniello, 2008). This work demonstrates that large scale exams have differential reliability for students whose English proficiency does not meet the linguistic demands of the test.

Martiniello (2008) and Abedi (2009) have proposed a number of directions in order to make standardized exams more reliable and unbiased for emergent bilinguals. One way to level the playing field for emergent bilingual students is to lower the linguistic complexity of test items (Abedi, 2009). Further challenges that tests present to emergent bilinguals are unfamiliar vocabulary, words with multiple meanings, and complex syntactical structures such as multiple clauses and long noun phrases (Martiniello, 2008: p. 356). Martiniello (2008: p. 363) suggests that in addition to content specialists, those familiar with the linguistic needs of emergent bilinguals should participate in the creation of test items that are linguistically understandable for these students so that they can demonstrate their knowledge on a particular item. Abedi (2009) goes further to call for an inter-disciplinary effort among psychometricians, linguists and teachers to determine how a particular test item is, or is not, related to language.

In addition to the linguistic demands of tests, there are a range of issues with how tests are administered and interpreted which can affect how bilingual students are assessed. The testing of English Language Learners is problematic with regards to each of the following steps: how students are classified, who administers the tests, and at what point students are tested (Abedi, 2004; Solano Flores, 2008). Solano-Flores (2008: 189) sums this up as “who is given tests in what language by whom, when and where.” One issue is the “English Language Learner” classification is not standardized across geographical areas and is not founded on a

common understanding of academic language proficiency (Abedi, 2004). Therefore, students classified as “English Language Learners” are far too heterogeneous in their linguistic ability to be considered a cohesive group. Each of these different components of testing – who is deemed a language learner, how does an assessment address the linguistic abilities of students, what are the conditions under which students are assessed, as well as their grade level and their language proficiency when are they assessed – adds to the complexity of the assessment for emergent bilingual students.

Given the difficulties in assessing emergent bilingual students, many researchers have focused on examining assessment practices and adapting them to obtain greater accuracy of emergent bilingual students’ achievement levels. One of the most utilized methods to do this is through testing *accommodations*. The following section provides a definition for testing accommodations, describes accommodation practices, and examines their potential in leveling the “assessment” playing field for emergent bilinguals. This discussion on accommodations is important to this study in that it provides background on how the assessment needs of emergent bilinguals can be met, and may provide insight into how their assessment data may be properly used by teachers.

### **Accommodations**

Accommodations are defined as modifications to the test or the test procedure to enable students to demonstrate their skills and abilities (Abedi, Hofstetter & Lord, 2004; Butler & Stevens, 2001). These modifications are not intended to provide students with undue advantage to perform higher than students who do not receive them. Rather they serve to decrease the disparity in scores between students who receive accommodations and others who have similar

levels of content knowledge (Abedi, Hofstetter & Lord, 2004; Butler & Stevens, 2001). For emergent bilinguals, accommodations are specifically targeted to “provide direct or indirect linguistic support” (Kim-Wolf, Griffin, Kao, Chang & Rivera, 2009).

Generally two types of accommodations (see Table 5) are used with emergent bilingual students (Butler & Stevens, 2004). One type of accommodations is *test modifications* which are designed to lessen the linguistic challenges of assessments. Accommodations of this type include test translations, but also consist of how the test is presented — such as the inclusion of visual supports and whether the language of the test is modified to lessen the linguistic load of an assessment. Other modifications that fall under this category also include the use of bilingual glossaries and/or dictionaries.

The second type of accommodation for emergent bilingual students is in adaptation of *testing procedures*. One type of change to the testing procedure is with regards to time: students are often given extra time or breaks during the assessment, or the test is broken up into sections which are administered over a longer period of time (Butler & Stevens, 2004). Another type of accommodation in testing procedure is varying the testing setting for students. Students can be tested in small groups or individually. This type of accommodation may also allow for the proctor, if bilingual, to deliver the test instructions in both languages. Finally, another change in testing procedure is to allow students to respond in their home language.

**Table 5. Frequent Types of Accommodations Implemented for Emergent Bilinguals**

<i>Test Modifications</i>	<i>Test Procedure</i>
<ul style="list-style-type: none"><li>• Test translations into home language</li><li>• Parallel/Test Transadaptations</li><li>• Modified language</li><li>• Inclusion of visual support</li><li>• Use of bilingual glossaries and/or dictionaries</li></ul>	<ul style="list-style-type: none"><li>• Extra time on test</li><li>• Varied setting (small group or individual)</li><li>• Breaks during the assessment</li><li>• Test is administered across a longer period of time in smaller chunks.</li><li>• Test instructions are read in both languages</li><li>• Student respond in home language</li></ul>

### **Test Translations**

Translating tests into students' home languages seems like a simple remedy to the problem of testing students in their emerging language. However, translations often are not equivalent to the original test. The properties of translated tests often change given that linguistic features and their difficulty vary between languages (Butler & Stevens, 2001). Therefore, translations of tests may broach the same content as the original test, but may be dissimilar psychometrically from the original version (García, 2010).

In practice, students are often provided with both translated and original versions of the test (García, 2010). They are expected to work in tandem between these two versions. In order for this procedure to work, students must have experience moving between test versions in the two languages. However, these skills are, for the most part, not taught in schools. Furthermore, considering the target population for the proposed study, students in Grades 3-5 (between the ages of 8 and 11), the need to move between two tests requires complex skills that these young test-takers may not have acquired or be developmentally ready for. It is questionable whether students this age can manage and manipulate the language information from both tests in order to lessen the influence of language on their testing performance.

Lastly, translations of tests into the home language may also miss their targeted goal because of the heterogeneity of the population for which they are intended (García, 2010). For example, although Spanish is the most common language spoken by emergent bilinguals, Spanish vocabulary and syntax varies from country to country and within regions. Therefore, although a student's home language may be Spanish, a translated version of a test may not reflect the language that large groups of students may commonly use.

### **Parallel Tests – Transadaptations**

Another practice to render assessments fairer is the development of parallel tests. Parallel tests are written in a student's dominant language but are not direct translations of tests.<sup>11</sup> In contrast to test translations, they are meant to bypass the issue of language as a confounder in the assessment of content because they are developed in the home language. In addition, parallel tests are normed for the bilingual population with which they will be used (García, 2010).

Parallel tests, however, have the same issue with regard to linguistic diversity as test translations do. Both test translations and parallel tests assume that students who take tests in their dominant language know the academic terms that appear on the test. This is often not the case, especially if content instruction for the given assessment has been in English.

### **Conceptual Scoring**

Conceptual scoring combines both a "testing accommodation" and a scoring practice. Conceptual scoring is premised on the fact that bilingual students' knowledge base overlaps two

---

<sup>11</sup> The term dominant language refers to the language in which a student has greater proficiency. However, this is a tricky term since students often straddle two linguistic worlds. They may be more proficient in one language in general, but have greater command of academic language in their second language. Therefore, language dominance is individual-specific and contextual.

different linguistic worlds (Bedore et al., 2005). In conceptual scoring students are provided with two versions of the test, one in English, the other in their home language. These tests cover similar tasks in the two different languages. This may be considered a testing accommodation. However, the way in which the assessment is scored is critical. Conceptual scoring represents an alternative method of testing which involves modifications to the test and its *interpretation*.

In order for conceptual scoring to be possible, students must be assessed in both of their languages. Conceptual scores match how student responses are distributed between their languages (Bedore et al., 2005). Bedore et al.'s (2005) study examined young bilingual children's vocabulary in both English and Spanish. The study revealed that bilingual students' conceptual score was similar to monolingual students' single-language score (Bedore et al., 2005). This finding confirms that single-language scores for emergent bilingual students, especially when given assessments in their emergent language, do not adequately represent the entire span of their knowledge base.

The descriptions of the accommodations above are offered to provide a landscape of the diverse ways in which summative assessments have been modified in order to adapt to the testing needs of emergent bilinguals. These examples are relevant to this research because they provide a background into the ways that emergent bilingual students are accommodated differently, which may impact teachers' buy-in and hence the knowledge they glean from the results of this type of assessment.

### **Summative Assessments: Problems & Directions**

Although this research does not aim to contribute the body of research having to do with the reliability and validity of summative assessments of emergent bilinguals, the discussion of

this type of research is critical since it highlights the difficult task of obtaining accurate student data for emergent bilinguals. At the current time no singular method or combination of methods has proven to be the silver bullet for making these assessments fair and accurate for all emergent bilinguals. The actual practice of how accommodations are implemented and what effect they have on English Language Learners' performance falls far from the intended results. First, accommodations for emergent bilinguals rest upon shaky empirical ground (Butler & Stevens, 2001). Although the most frequent types of accommodations provided to emergent bilingual students are extended time, translations and bilingual dictionaries, none of these assessments have been empirically demonstrated to be effective methods (Abedi, Hofstetter & Lord, 2004). Furthermore, when accommodations are randomly assigned to both emergent bilinguals and English language proficient students, some research has demonstrated that English proficient students perform at much higher levels than emergent bilinguals (Abedi, Hofstetter & Lord, 2004). This finding invalidates the accommodation, given that the purpose of the accommodation is to solely affect the performance of the emergent bilingual students.

One of the reasons why accommodations may not work for emergent bilingual students can be attributed to how they are implemented. It has been suggested that accommodation has been largely ineffective because they do not address these students' linguistic needs appropriately (Abedi, 2009). For example, a student may be offered extra time on a test, but this accommodation will not lessen the linguistic demand of the test items. Therefore, in most cases, accommodations that are provided to emergent bilinguals are often not related to a student's specific linguistic need.

However, several promising avenues for accommodations for emergent bilinguals have been reported. Abedi, Hofstetter & Lord (2004) find that student performance was directly related to the language of instruction. For example, bilingual students instructed in English performed better on an English version of a test, as compared to a Spanish version covering the same content. This suggests that the language of assessment should match the language of instruction. In addition, accommodations may have different effects depending on the language proficiency of the emergent bilingual (Butler & Stevens, 2001). Given that bilingualism exists on a continuum, emergent bilinguals at the earlier end of the proficiency continuum may require different accommodations than those students at later stages of proficiency. In contrast, the modification of test language and the use of “customized” dictionaries and glossaries, rather than general ones, have proven to be effective in better isolating the content to be assessed from the language in which it is presented (Abedi, Hofstetter & Lord, 2004). Although these accommodations provide exciting opportunities to render fairer assessments for emergent bilinguals, they are rarely implemented. In a study which analyzed all 50 states’ accommodation policies, the most common accommodations found were “extended time, frequent breaks, individual testing, small group testing, and testing in a separate location (Kao et al., 2009).”

Butler & Stevens (2001: 421) note that, “currently no single assessment approach has proven to be an effective solution to the problem of measuring ELL content knowledge in English when student English language ability is weak.” Despite this when accommodations are thoughtful and appropriately administered, emergent bilingual students do benefit (Kopriva, 2007). This point of appropriately assigning and administering assessments to match student needs cannot be overlooked because when emergent bilingual students receive *inappropriate*

accommodations, their performance is unaffected and therefore accommodations are ineffective (Kopriva, 2007).

Current testing practices fall far from the ideal for accurately assessing emergent bilingual students, yet they are still expected to drive instruction. Despite the lack of reliability of large scale exams for emergent bilinguals and the limited validity of the inferences made about these same learners from their results on standardized tests, school administrators and teachers are often required to use this data to fuel curricular decision-making. The majority of the literature on summative assessment and bilingual students considers the effectiveness of practices aimed at making assessments fairer for emergent bilinguals, rather than *how assessment is used* to steer instruction. The ways in which school leaders and teachers navigate the use of data from large scale exams is largely unexplored. This critical step of using this data appropriately is essential if student learning is to be achieved (Heritage & Chen, 2005). The next section examines the research conducted on formative assessments as they are linked to student learning.

### **Formative Assessment & Bilingualism**

Formative assessment is closely tied to student learning (Black et al., 2003; Falk et al, 2007; McCutchen et al, 2009). An assessment is considered formative based on its *use* to inform teacher practice, rather than its format (e.g. multiple-choice versus extended response). Black et al. (2003: 122) argue that formative assessment is a *process*. As previously described, Popham contends that formative assessments must be used as the first step in a learning progression (2008). As such they catalyze the process of informed decision-making for teachers about their students. Although formative assessments may encompass a wide range of

assessment formats, including tests, rubrics, checklists, various methods of documenting reading, portfolio tasks, and observations, these types of assessments become formative once teachers embed them within the teaching practices. Because of its embedded nature, formative assessment may catalyze changes to teacher practice (Black et al., 2003). Teachers who are involved in formative assessment may reconfigure classroom practices based on their assessment of students and the corresponding teaching practices which emerge from these assessments.

Black & William's (1998) groundbreaking review of studies of formative assessments concluded that these types of assessments have the potential to "produce significant and often substantial learning gains." In a more recent study, Heritage et al. (2009) approached the question of formative assessment from the perspective of teacher knowledge. They found that although teachers were able to identify students' abilities and skills, they were less able to formulate next instructional steps (Heritage et al., 2009). This suggests that although teachers do learn from formative assessment, it is questionable if they are able to translate this knowledge into action. Possible explanations from this shortcoming may have to do with the lack of professional development about using assessment and support of how to forge the path between teacher knowledge and instructional practices.

Another impediment to realizing the full potential of formative assessments may be due to how teachers perceive and engage in such assessments. One reason that the full potential of formative assessment may not be realized is because of a spill-over effect from how teachers experience summative assessments. Black et al. (2003: 54) noted that, "frequent summative testing dulled the message about the means to improve, replacing it with information about

successes or failures.” The formative assessments that teachers administer are often prescribed to them; thereby rendering them less useful for the purpose of improving student instruction.

Additionally, formative assessments may also follow an “assessment script.”<sup>12</sup> The strict adherence to an assessment script may not allow teachers to make adaptations that would provide them with insight into students’ learning. This has particular significance for teachers who work with emergent bilingual students. Pre-packaged formative assessments are intended to assess students in their home language. Therefore, unless teachers of emergent bilingual students adapt pre-package assessments to account for bilingualism, the use of these assessments may also result in inaccurate data concerning emergent bilingual students.<sup>13</sup>

Although formative assessments are deemed more appropriate to use with emergent bilinguals, how these types of assessments can be adapted to better assess these students is not fleshed out by those who advocate them. The effort to make assessment fairer for emergent bilingual students has mostly been limited to the realm of summative assessments. In general, there is a very little literature on how formative assessments can be rendered fairer for emergent bilinguals. More often than not, teachers’ formative assessments of emergent bilinguals are not

---

<sup>12</sup> Some formative assessments, such as formal running records, may come in the form of pre-packaged assessments. These assessments are devised to follow a script – providing teachers with prompts and questions to pose to students. Teachers may be instructed by school administration to follow the script without deviation, thus adding to the view that formative assessments are tied to school mandates rather than the pursuit of what teachers deem relevant knowledge of students.

<sup>13</sup> A small pilot study which involved observations of 2 teachers and a total 4 running records that I conducted (May 2009) revealed that teachers of emergent bilingual students made adaptations to pre-packaged running records by asking additional questions and questions in both of the student’s languages to probe student comprehension, expressive language abilities and reader self-confidence – all important ingredients in a student’s emerging identity as a reader in a second language.

guided by a theory of bilingualism (August, 2006; Bachman, 2001). However, two bodies of knowledge – descriptive process and dynamic assessment – may provide some promise for charting a course of how formative assessment can be better aligned to the educational goals of bilingual students.

Descriptive processes developed from the work of teachers at the Prospect School in Vermont (Carini, 2000). These processes allow teachers to generate knowledge about students and school through the intense study of students, teachers and their work. Descriptive processes, “focus on what is visible and available for sensory description” (Ascenzi-Moreno et al., 2008). Although not considered a type of assessment, descriptive reviews of a student or their work have the potential to impact assessment because it provides holistic information on the students as a learner over time.

Ascenzi-Moreno et al. (2008) described the work of a collaborative group of teachers and university professors who conducted several longitudinal case studies of bilingual children using descriptive processes. One case study illustrated that students’ bilingualism as contextual and extending beyond classroom-based language and skills. Ascenzi-Moreno et al. (2008: 398) stated, “when we think of Manny’s language as also encompassing his use of it, how he expresses and reveals himself to others, it reminds us that children learn and create because they are participants in the world, a much larger context than the words of our classrooms.” Often student language use is viewed in relation to the language that is prioritized in school. The insight that student language use exceeds “school language” directed teachers to broaden the repertoire of language used in classroom. This insight into language allowed students with different levels and modes of using language to be brought into the classroom community.

Although the research does not specify any particular assessments or assessment practices, it does refer to the importance of descriptive processes as the basis for pedagogical decision-making and as a means to enhance teachers' working knowledge of language and how students experience language.

Finally, work on dynamic assessment also offers a view on how assessments can incorporate a holistic understanding of students' language use. Dynamic assessment is the introduction of questions and/or tasks attuned to a student's demonstration of knowledge. In this method of assessment, the line between instruction and assessment is blurred. Furthermore, what is being assessed is not only the level of knowledge that the student possesses, but also how a student approaches a task, how much the student is able to learn within a period of time and how a student applies knowledge to a new situation. As Poehner (2007) notes:

In lieu of understanding assessment as the observation and recording of individuals' behaviors for the purpose of inferring underlying abilities, assessment in the dynamic sense involves the transformation of these abilities through dialogic collaboration between learners and assessor-teachers or mediators. (p. 423).

Dynamic assessment is based on the interactive nature of cognitive development as posited by Vygotsky (Gutiérrez-Clellen & Peña, 2001: 212). In addition, dynamic assessment departs from traditional assessment practices in that rather than solely documenting student performance, it aims to understand the scope of student learning through interaction between teacher and student (Poehner, 2007). Therefore, the aim of dynamic assessment is not solely to assess the academic content a student possesses, but also to determine the range at which the student performs.

This method of understanding student learning is similar to the descriptive processes in that it does not isolate students from contexts that deeply influence language use, but rather embraces them. The work described above – descriptive processes and dynamic assessments – are both rooted in context and process. Each of these aim to provide a snapshot of students as they grow as learners. Therefore, both of these methods are cumulative and on-going, in contrast to typical standardized assessments. Furthermore, they are not entrenched in fixed procedures, but rather are based upon evolving teacher knowledge of context and students. Although neither overtly claim affiliation with a particular view of a theory of language learning, they are aligned with a theory of language acquisition that is contextual. There is untapped potential in both of these methods to assess emergent bilinguals. Yet, these methods are not commonly used by teachers of emergent bilinguals. In fact, there is evidence to suggest that professionals working with bilingual students are more likely to turn to assessments designed for the general student population, rather than those adapted for use with bilingual students (Caesar & Kohler, 2007).<sup>14</sup>

The relationship between formative assessment and student learning has been demonstrated empirically (Black et al., 2003; Heritage et al., 2009). Nevertheless, many questions remain, especially with respect to how formative assessments impact practice for emergent bilingual students. Formative assessment by definition is rooted in classroom practice. In the case of emergent bilingual students, it must also be responsive to the students' linguistic characteristics and to a teachers' or school's conceptualization of bilingualism. Yet, important questions concerning what knowledge teachers gain from formative assessments of emergent

---

<sup>14</sup> The study cited, *The state of school-based bilingual assessment: Actual practice versus recommended guidelines*, by Caesar & Kohler (2007) was focused on the use of assessments by Speech-Language Pathologists (SLP) when assessing bilingual students.

bilingual students, what data they find relevant to teaching, and how they translate what they find out about these students into practice must be addressed. In the next section, I offer a definition of reading and its components, as well as describe the research base on the reading development of emergent bilingual students.

#### **Part 4: Reading Development of Emergent Bilinguals**

This thesis examines how teachers learn from reading assessments of emergent bilinguals. Reading plays foundational role in all areas of academic achievement. Therefore, it is important to describe the ways reading develops and varies for emergent bilinguals and monolingual students.

The ability to read proficiently is at the heart of school success (Bialystok, Luk & Kwan, 2005; Proctor, August, Carlo & Snow, 2005). Reading is a critical process that involves decoding, fluency and comprehension (Cooper, Kiger, Robinson & Slansky, 2010). Students become competent readers by simultaneously engaging their knowledge about sounds, words, and sentences in order to make meaning of a given text. In fact, Harvey & Goudvis (2007) contend that the ultimate goal of reading is comprehension. They state that reading is “thinking” and is constructed while students make personal connections and interpretations about what they read. They write, “comprehension means that readers think not only about what they are reading but about what they are learning” (Harvey & Goudvis, 2007: 15). This emphasis on comprehension stands in contrast to what most people consider as competent reading at the elementary school level. Reading at this level is often regarded sufficient once decoding skills, fluency, and phrasing are adequate. Although these technical aspects of reading are important, the acquisition of these skills do not automatically lead to comprehension.

Decoding is based on students' ability to employ their knowledge of phonemic awareness to make sense of written words (Cooper et al., 2010). Phonemic awareness is the ability to know and manipulate phoneme, the smallest unit of sound in a language. For example, "sh" is a sound that is associated with the two letters "s" and "h" combined. Fluency and phrasing refer to students' ability to know words automatically and accurately in order to read at a pace that facilitates comprehension (Cooper et al., 2010). In addition, comprehension skills such as determining importance, predicting, inferencing, questioning visualizing and synthesizing are essential in order for students to make meaning of a given text (Harvey & Goudvis, 2007). It is important to note that all of these skills and strategies taken together lead to the development of a competent reader. While students who can solely decode and read fluently at the surface appear to "read," they are not truly readers in the sense that they are engaging in the construction of meaning.

Pacheco (2010: 314) maintains that for emergent bilinguals in particular, it is important to expand the commonly-held notion of reading which is capped at decoding and literal comprehension to include, "sophisticated, deep, and critical reading practices that are associated with higher academic learning and achievement." In particular, reading for emergent bilinguals is often considered adequate if these students acquire skills such as decoding and fluency, critical skills which facilitate comprehension, but do not ensure it. In addition to these skills, students must also possess language skills. These include an understanding of vocabulary, grammatical structures (syntax), knowledge of text features and the ability to connect their background knowledge to what they read (Snow, 2008).

Understanding how the process of learning to read compares for monolingual and emergent bilingual students is crucial for the purpose of meaningful assessment use and adequate instruction. Differences in how reading develops between these two types of students can inform how assessments must be analyzed differently for each group and how instructional practices can best be suited to enhance students' bilingual reading development.

There is a wealth of research that shows that monolingual readers benefit from understanding letter symbols and sounds, vocabulary knowledge, oral language proficiency, and world knowledge (Snow, 2008; Bialystok, 2007). In this vein, learning to read for both monolingual and bilingual students shares some similarities. For both populations, oral language proficiency, “general vocabulary language,” and listening comprehension are skills that are positively linked to successful reading (Proctor, Carlo, August & Snow, 2005: 247). However, the paths for acquiring these early predictors of reading success may differ for emergent bilingual students. Bialystok (2007) underscores that the skills necessary for students to learn to read – in particular letter symbols and sounds, oral language, and metalinguistic skills – are influenced by student bilingualism. In another paper, Bialystok, Luk & Kwan (2005: 247) maintain that emergent bilingual students develop background literacy skills differently than monolingual students. For example, some emergent bilingual students may be able to transfer knowledge of reading skills between their home language and their second language (Proctor et al., 2010; Bialystok, 2007; Cummins, 1979). These differences are key to understanding how instructional decisions in the teaching of reading should differ for emergent bilingual students.

Moreover, within the emergent bilingual population, there are differences between those emergent bilingual students who learn to read solely in their additional language or in two

languages simultaneously. It is important to consider how reading develops for emergent bilingual students in programs with different instructional approaches and various levels of support in their home language. In this study, all teachers interviewed worked with emergent bilinguals who are taught to read in two languages (Spanish and English); teachers who were administered the survey, worked with both types of emergent bilinguals, those who are learning to read in English only and those that are learning to read in two languages.

For those emergent bilingual students who are learning in two languages, the transfer of literacy skills between languages is critical (Cummins, 1979). This ability to use literacy skills developed in one language to support literacy in another is referred to as transfer or interdependence. In a study by Proctor et al, (2005: 53) bilingual readers drew on additional comprehension strategies, notably translation and cognate awareness, to further enhance their comprehension outcomes. Baker also notes that skills such as decoding and reading strategies are generalizable and can transfer from a students' repertoire of reading abilities in the home language to the second language (Baker, 2011: 322). Transfer is a positive effect of bilingualism – it allows students to enrich their language development in two (or more) languages using their knowledge acquired through one. Bialystok echoes the benefits of transfer but warns that it may not occur on its own. She writes, “children who have learned skills in one language can potentially benefit from that mastery by applying them to another. Even though such transfer is not automatic nor assured, when it does happen, the consequences are always salutary (2007: 71).”

In fact, although research provides evidence of linguistic transfer, how this transfer occurs within a classroom setting has not been the focus of these studies (Baker, 2011: 323).

Proctor et al. (2005: 18) note that “to reap the potential benefits of interdependence, instruction must focus on language and literacy development in both languages, and indeed focus directly on explicitly teaching for transfer.” Therefore, for emergent bilinguals who are learning to read in two languages, the transfer between reading skills acquired in different languages may need to be explicit in order to facilitate this benefit.

Additionally, there has been documentation that students who learn to read in both their home and second languages secure academic benefits (Slavin & Cheung, 2005). One way in which they do is that these students tend to outperform monolingual students because they develop early an understanding that words are symbolic and convey meaning (Bialystok, 2007). Moreover, as Baker(2011) maintains there is value in students knowing how to read in more than one language and thereby able to access different perspectives communicated through different languages.

In contrast, for emergent bilinguals who do not have the opportunity to develop reading skills in their home language, listening skills and vocabulary knowledge in their additional language are important predictors of reading ability (Proctor et al, 2005, p.). Proctor et al., posit that although these general skills – listening and vocabulary knowledge – are important both to monolingual and emergent bilingual students who are learning to read, the ways in which they may acquire these skills may be very different (2005). In fact, in their paper on emergent bilinguals’ L2 (second language) reading, these researchers bring up the beginner’s paradox. How do emergent bilinguals acquire high levels of listening comprehension and vocabulary, critical components of successful reading, in the second language being learned, if their proficiency level in that language to start off with is low? Therein lies the importance to

instructional practices for emergent bilinguals that differ from those practiced with monolingual English students. Although emergent bilingual students may need similar skills to monolingual students in order to be successful readers, they need different instructional paths to get there.

Learning to read for emergent bilinguals may take various routes, as emergent bilinguals are a diverse group and are educated through different approaches. Although similar underlying skills such as listening comprehension, vocabulary, oral language proficiency, and metacognitive skills are important for both monolingual and emergent bilingual reading development, the way in which emergent bilingual students acquire these are different. The instructional needs of emergent bilingual students in reading vary depending on the stage of language acquisition of students, but differ from monolinguals in that oral language experiences must include academic and culturally-specific language structures, language-specific knowledge of letter-sound relationships and words, as well as the ability to know how to transfer skills and strategies between languages (Bialystok, 2007). In sum, teachers need to employ different strategies in order to teach emergent bilingual students to be successful readers than when teaching monolingual students to read.

Correspondingly, the assessment of emergent bilingual students' reading must take into account that emergent bilingual students take a different route from monolingual students in learning to read. For example, in order to understand how emergent bilingual students' reading develops it is important to consider these students' proficiency levels in both languages, reading levels in both languages, oral language use, vocabulary knowledge, use of metacognitive strategies, and background knowledge among others. Yet, as Escamilla (2006) describes teachers generally lack knowledge about how the process of acquiring literacy between emergent

bilinguals and monolingual students differs. This lack of knowledge impacts how teachers administer assessments to emergent bilinguals and ultimately use resulting data to effectively teach students who are learning to read *bilingually*.

## **Summary**

Assessment is increasingly viewed by policymakers and educators as a way to close the achievement gap. This means that emergent bilinguals are included in yearly testing one year after their arrival. The majority of research on the impact of summative assessments on emergent bilingual students is from the perspective of making these more valid and reliable, rather than examining how they are *used* by educators as well as their capacity to *develop useable knowledge* to impact student learning. In contrast, the available research on formative assessment provides a more complete picture of how these various types of assessments impact teacher knowledge as well as student achievement. However, this literature has not considered how teachers make sense of the formative assessment data of emergent bilingual students.

Lastly, research on the reading development of emergent bilinguals while reporting some similarities to that of monolingual students' reading development, highlights that bilingualism does impact the way literacy is acquired by these students. As such, emergent bilingual student reading develops differently. Therefore, assessments of emergent bilingual students must take these differences into consideration. Teachers must also know how the instruction of emergent bilingual students in reading differs from the instruction of monolingual reading in order to adapt assessments to reflect these differences.

This doctoral research is situated within these gaps in these bodies of literature, as it attempts to explore what knowledge teachers develop from assessments of bilingual students and

how they craft instructional practices based on this knowledge. The study does not pretend to provide information about how to change the current assessments employed by teachers, but rather to understand what teachers do with existing reading assessments of emergent bilinguals in order to shape appropriate instruction for these students.

## **CHAPTER 3:**

### **Methodology**

#### **Introduction: A Mixed-Methods Study**

This research used data from interviews and a survey in order to explore the question of what types of knowledge teachers gain from the use of assessments of emergent bilinguals and how this knowledge relates to instructional practices. The practice of combining qualitative and quantitative practices in order to study a phenomenon is referred to as mixed-methods. Johnson & Onwuegbuzie (2004: 17) defines mixed-methods research as “the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language in a single study.”

Since this research involved the design, development, implementation, and analysis of two complementary sets of data, for the sake of clarity, this section will be divided into six sections. The first section on research design details the alignment of research questions to data, methodologies and, in the case of the quantitative research, hypotheses. The second section describes the tandem development of the interview protocol and survey that were designed for this study. The third part presents the assessments that either were included in the survey or that the respondents in interviews referred to. The descriptions of the assessments are featured here in order to complement the description of the research instruments by providing readers with the assessment context in schools. Section four describes the research sites and sampling methods for both the interviews and survey. Section five describes how the analysis was conducted for both methods. Lastly, section six describes the limitations and validity concerns for the research.

## **Part 1: Research Design**

Mixed-methods were used during the development and analytical stages of the research. This approach is particularly suited to this study's research questions because these two methodologies complement each other to provide a portrait of the complex processes of developing teacher knowledge and instructional practice from the assessment of emergent bilingual students. A qualitative study alone would provide in-depth narratives highlighting what teachers learn from assessments and the instructional practices they develop as a result. A quantitative study provides trends which may be generalizable. Therefore, through the integration of these two methodologies, multiple perspectives regarding how the assessment-instructional cycle works in practice for emergent bilingual students is offered.

The following research question and methodology matrix displays the research questions examined in this mixed-methods study. By setting the rationale, data, methodology, and quantitative data analysis alongside the research questions, the matrix offers readers a complete overview how each research question was studied as well as how the interview and survey methods complement each other.

**Table 6. Research Question & Methodology Matrix**

Research Question	Rationale	Data	Methodology & Hypotheses
<p>1. What kinds of knowledge do teachers of emergent bilinguals gain from examination of reading assessment data of emergent bilinguals?</p> <ul style="list-style-type: none"> <li>▪ How does the knowledge that teachers gain differ when using large-scale standardized assessments and using classroom-based assessments?</li> </ul>	<p>To describe what teachers learn from the assessments they currently use with emergent bilingual students.</p> <p>To examine self-reported changes in knowledge as a result of using a large-scale standardized assessment (ELA)<sup>15</sup> or a classroom based assessment (running record).</p> <p>To compare how knowledge acquired from the assessment of emergent bilingual students differ between types of assessments (ELA and running records).</p>	<p><i>Interviews:</i> Teacher narratives about what teachers learn from either ELA or running records data about their emergent bilingual students</p> <p><i>Survey:</i> Teacher self-reports about changes in knowledge after using ELA or running records data of emergent bilinguals.</p>	<p><i>Interviews:</i> Identify codes about teacher knowledge derived from the interviews.</p> <p><i>Survey:</i> <u>Hypotheses:</u> 1. Teachers do report changes in knowledge from assessment use of emergent bilinguals. 2. Teachers report greater changes in knowledge from use of formative data of emergent bilinguals than from use of summative data.</p> <p style="text-align: center;"><b>Data Analysis</b></p> <p><i>Interviews:</i> Develop and describe themes about teacher knowledge and learning emerging from codes.</p> <p><i>Survey:</i> Teacher self-reports of changes in knowledge from analysis of ELA or running record data displayed by percentile distribution.</p> <p>Teacher self-reports of changes in knowledge from analysis of ELA or running record data analyzed with Wilcoxon signed rank test (non-parametric t-test).</p> <p>Correlations of teacher self-reports of knowledge gained with frequency of assessment use.</p>

<sup>15</sup> ELA stands for English Language Arts exam, given yearly by New York State.

Research Question	Rationale	Data	Methodology & Hypotheses
<p>2. What changes, if any, in reading instructional practices for emergent bilinguals result from teachers of emergent bilinguals' use of large-scale standardized assessments and/or classroom-based assessments?</p>	<p>To examine self-reported changes in instructional practices for emergent bilinguals as a result of either using ELA or running records data.</p>	<p><i>Survey:</i> Teacher self-reports about frequency of assessment use (ELA and running records).</p> <p>Teacher self-reports of changes in instructional practices after viewing ELA or running record data of emergent bilinguals.</p>	<p><i>Survey:</i> <u>Hypotheses:</u> 1. Teachers differ in the implementation of instructional practices as a result of either ELA or running record use.</p> <hr/> <p><b>Quantitative Data Analysis</b></p> <p>Teacher self-reports of changes in instructional practices after viewing either the ELA or running records for emergent bilinguals displayed by percentile distribution</p> <p>Teacher self-reports of changes in instructional practices resulting from examination of either ELA or running record data analyzed with Wilcoxon signed rank test.</p>
<p>3. What changes, if any, in reading instructional practices for emergent bilinguals result from the types of knowledge gained by teachers of emergent bilinguals from large-scale standardized assessments and/or classroom-based assessments?</p>	<p>To inquire how changes in instructional practices for emergent bilinguals are associated with types of teacher knowledge acquired through either ELA or running record data.</p>	<p><i>Interviews:</i> Teacher narratives about what instructional practices they implement after examining either ELA or running record data.</p> <p><i>Survey:</i> Teacher self-reports of changes in instructional practices after viewing ELA or running record data of emergent bilinguals.</p> <p>Teacher self-reports about changes in knowledge after using the ELA or running records data of emergent bilinguals.</p>	<p><i>Interviews:</i> Identify codes about instructional practices derived from the interviews.</p> <p><i>Survey:</i> <u>Hypotheses:</u> 1. Teachers report greater changes in instructional practices knowledge gained from running records. 2. Teachers' use of instructional practices is positively associated with greater levels of teacher knowledge.</p> <hr/> <p><b>Data Analysis</b></p> <p><i>Interviews:</i> Develop and describe themes about instructional practices emerging from codes.</p> <p><i>Survey:</i> Correlations between instructional practices and types of teacher knowledge acquired through either ELA or running records.</p>

The first research question asks what kinds of knowledge do teachers gain from looking at emergent bilinguals' assessment data. Also, part of this question is to compare the type of knowledge gained from large-scale standardized assessments with classroom-based ones. The data that was collected to answer this question stemmed from both interviews and the survey. In particular, the qualitative research questions aimed to glean teachers' impressions about what they learned about emergent bilinguals from these assessments. The questions in the interview protocol were "open," thus inviting any type of response that teachers would like to provide. While the survey collected similar information, respondents were offered "fixed" choices about the types of knowledge that they acquired, as will be described in a later section. Furthermore, the survey provided respondents with degrees of perceived learning. Therefore, the last part of this question will be answered solely with quantitative data.

The second and third questions that drive this doctoral thesis deal with instructional practices implemented after examining assessment data of emergent bilinguals. Specifically these questions examine how instructional practices are related to use of assessments and teacher knowledge. The second question relies solely upon data from the survey. Through analysis of frequency distributions, the frequency with which teachers report implementing instructional practices upon consulting either English Language Arts Exam (ELA) or running record data of emergent bilingual students is explored.

The last question utilizes both interview and survey data. The interviews provide examples and opinions from teachers about what instructional practices they implement as a result of what they learn from assessments. Data from the interviews have the potential to provide teachers with an open forum in which to talk about what they do after they examine their

data about emergent bilingual students. On the other hand, survey data again fixes the type of instructional practices that they are offered — thus capturing teacher responses about a standard set of instructional practices. The survey data can be analyzed to understand the extent of the relationship between teacher knowledge and instructional practices.

## **Part 2: Instrumentation**

The data for this research was gathered from an *interview protocol* and a *survey* designed for this research. The choice of these two methods for data collection stems from the abilities of each of these methodologies to capture data about the same phenomenon, but of a different quality. In a paper reviewing research methodologies, Desimone (2009: 189) notes that, “well-constructed and administered observation, interviews, and surveys can elicit much the same information.” In other words, each of these methods may be related to each other, but because of their unique qualities, may uncover varied perspectives on the same topic. Therefore, although the interview and survey collected information that overlapped, the quality of the information was different from each other.

Both the interview and the survey designed for this study inquire about teachers’ use of *reading assessments*. The reason why reading assessments were chosen for this study is because of the important place that reading has in schooling. In addition, the practice of administering reading assessments is more widespread among teachers than is the practice of administering other types of literacy assessments. Furthermore, teachers are instructed through their professional education and professional development experiences how to interpret reading assessments, both large-scaled standardized and classroom-based, and often lack the skills to examine students’ growth areas of literacy such as writing or speaking.

The assessments that are focused on in this study are examples of the summative and formative assessments that are most commonly administered at schools and used by teachers in New York City. These assessments, to be described in detail later in this methodology section, are similar to reading assessments administered throughout the country. However, in schools and classrooms they are not usually referred to as either formative or summative assessments. Rather, they are commonly called either “large-scale standardized” or “classroom-based assessments.” Therefore for the purpose of the field research, I will shift from using the terms summative and formative assessment, which were used in literature review, to the terms, “large-scale standardized” designating summative assessments and “classroom-based assessments” for formative assessments.

Also, the term “English Language Learner” is used in both the interview and the survey for the purpose of being understood by interview and survey participants. This decision to use this term for the fieldwork was made because of its widespread use by teachers. Therefore, although the term emergent bilingual is used throughout the literature review and analysis, English Language Learner or “ELL” used in both the interview questions and for the sake of clarity with respondents.

### **Timeline**

I took a partially sequential approach to designing and collecting the data. The design of the interview questions and survey began in May 2010. Interview questions were initially tested out in July 2010 with two graduate students. As will be described further, interview questions were further refined until the qualitative data collection began in early November 2010.

Interviews were conducted through May 2011. The original goal was to collect 24 interviews; a total of 20 interviews were conducted.

The data that emerged from the interviews served partially to refine and enhance the constructs used in the survey, specifically those constructs dealing with teacher knowledge. The survey was also refined through a “cognitive interview,” also to be described later, and launched in April 2011. Data collection for surveys was closed in late June 2011. Although 100 collected surveys was originally the goal, 63 were collected by the time the survey web site was closed.

## **Interviews**

Interviews enable researchers to collect detailed data about a particular phenomenon from a personal perspective (Codó, 2008). In addition, interviews can be considered behavioral and linguistic events representing social conventions and discourse between interviewed and interviewer (Mishler, 1986: 10). However, in this research, I am primarily concerned with content emerging from the interviews and will analysis them solely for content. According to Codó, content analysis of interviews can offer both facts about the respondent as well as their personal perspective (2008: 161). The interview protocol was constructed to gather both types of data (see Appendix 1).

The interviews collected information about the types of reading assessments teachers used for emergent bilinguals and how these contributed to their knowledge and instruction. In addition, their opinions about these assessments were also collected. The interviews conducted for this study followed an interview protocol that contained two types of questions: standardized and informal. The first type of question was standardized. These questions were constructed with the help of Margaret Heritage (telephone conversation, 3/17/10). Standardized questions

were posed to all respondents. The purpose of the set of standardized questions was to ensure that all teachers responded to the same inquiries about assessments without veering off track during the short time that was available for the interview. These questions asked teachers to describe the purpose of a given assessment, why they use it, what they find useful about it, an example of how it provides knowledge about language development, and lastly, how it leads to instructional steps for emergent bilingual students. These questions aim to describe the processes by which teachers use and learn from the reading assessments of emergent bilinguals.

The second type of questioning was “informal” or unstructured and emerged from the particulars of the interview. Data-wise, unstructured interviews are considered to gather more nuanced data with detailed that are “more varied and numerous” (Codó, 2008: 165). The objective of these questions was to dig deeper, to clarify points made in the interview, and ultimately to gather the underlying story of teachers’ developing knowledge and engagement in the use of assessment to improve instruction for their emergent bilingual students. Furthermore, the objective of these questions was to provide a glimpse into the institutional, social, and personal factors which influence the learning from and use of assessments of emergent bilinguals. Lastly, questions were posed at the end to all respondents about their background characteristics. These included information about the grade they taught, how many years they taught for, their certification, and if they held a bilingual extension or not, among others. The data collected from these questions is factual and will be used to describe the sample of teachers interviewed.

The interview protocol intentionally allowed for both types of interview questions, standardized and unstructured so that the interviews gathered systematic information as well as

the varied meanings that respondents make of their use of assessments with emergent bilingual students in classrooms. The collection of data from the standardized questions allows for the comparison of responses between respondents. The unstructured portion of the interview allows for the gathering of particulars of each respondent's students, classroom, school and personal history and vision of assessment. Both types of data are critical for understanding what teachers of emergent bilinguals learn from assessments and subsequently implement.

### **Surveys**

Surveys have been one of the primary methods that researchers have used to measure the implementation of large-scale policies (Desimone, 2006: 641). Surveys are often the preferred method of data collection for evaluation purposes because they are efficient to administer and can potentially gather data systematically from a large number of respondents (Codó, 2008; Picciano, 2004). Furthermore, and once field tested, surveys yield valid results that are highly correlated with classroom observations (Desimone, 2006: 642).

Surveys work particularly well to gather information, such as frequency and self-report data on behaviors, opinions, and knowledge (Fowler, 1995). One of the qualities of surveys is that their format and content is standardized. This standardization allows for comparisons across respondents (Fowler, 1995: 2).

Careful attention to survey construction is necessary to develop a reliable and valid instrument. Fowler (1995: 2) identifies the following essential criteria in developing a strong survey: 1) questions are consistently understood, 2) questions are administered consistently, 3) questions elicit the same type of response from all respondents, 4) questions are answerable, and 5) respondents are "willing to give correct and valid answers."

The construction of the survey took into consideration the concerns listed above (see Appendix 2). For example, in the survey designed for this study, definitions are provided when necessary, so that respondents refer to this definition as the basis for their answer (Fowler, 1995, 85). Furthermore, the repetitive quality of the questions and answer format in the survey allowed respondents to focus on the content of the question rather than how it was posed. The initial instructions and formatting used in this survey was modeled after, “Teacher Questionnaire – Year 4” used for the Study of Instructional Improvement, a large scale study conducted at University of Michigan.

Recent educational research has made notable strides in the development of surveys that fulfill the criteria listed above. Desimone (2004) describes the value of cognitive interviews in developing strong survey instruments. Cognitive interviews with respondents ensure that questions measure what they intend to. According to Desimone (2004: 6), “the cognitive interview is a method that does allow for in-depth analysis of individual items.” During a cognitive interview the survey is piloted individually before being administered to a final sample. It is a method of revising and editing the survey so that the content is clear and understandable to respondents. Cognitive interviews provide detailed feedback about what respondents understand questions to mean, as well as ask them to give suggestions regarding wording and formatting. This process of conducting cognitive interviews is a vital tool in confirming that survey questions measure what they set out to or point to the need of revision of a particular item.

One cognitive interview with a dual language bilingual grade four teacher was conducted in February 2011. Primarily changes to the formatting resulted from the cognitive interview. Furthermore, the survey was vetted by another graduate student who specializes in survey

research. Recommended edits following his comments were made. These edits included simplifying the language of the instructions preceding each section of questions and streamlining the instructional practices portion of the survey. Another teacher took the revised survey in order to ascertain the time that it would take to fill out.

The survey for this study is comprised of five sections. Each part is designed to gather distinct information:

- 1) frequency of assessment use
- 2) extent of knowledge gained by teachers
- 3) frequency of implementation of instructional practices for emergent bilingual students
- 4) teacher background and characteristics
- 5) open response for written opinions.

From these different sections, five different types of variables were collected from the survey. These include:

- 1) frequency of use of assessments
- 2) type and extent of knowledge acquired
- 3) frequency of instructional practice used as a result of analysis of ELA data of emergent bilingual students
- 4) frequency of instructional practice used as a result of analysis of running record (RR) data of emergent bilingual students
- 5) demographic information.

Table 7 provides an overview detailing the variables relevant to the analysis. For a complete list of variables captured by the survey see Appendix 3. For frequency of use, there are two variables: one for the number of times teachers report using the ELA data of emergent bilingual students and one for the number of times teachers report using running record data of emergent bilingual students. For the second type of variable – the type and extent of knowledge acquired – in the literature review I specified four types of teacher knowledge that are examined

in this research (see Table 4). For each of these types of knowledge – pedagogical content, student, assessment, and bilingual knowledge – there are two variables: one which represents, for example, pedagogical content knowledge that stems from viewing ELA data of emergent bilinguals and another that emerges from analysis of running record data of emergent bilinguals. Thus, there are a total of eight variables for teacher knowledge. Similarly, for each of the instructional practices specified in the survey, there are two variables: one which represents the frequency of implementation of a given variable after consulting ELA data of emergent bilinguals and another which represents the implementation of the same variable after consulting running record data of emergent bilinguals. Thus, there are two sets of variables for instructional practice: 17 instructional practices which stem from examination of ELA data of emergent bilinguals and 17 instructional practices that come from looking at running record data of emergent bilinguals. Lastly, each respondent answered a number of demographic questions.

**Table 7. Overview of Variables used in Quantitative Analyses**

Type of Variable	Variables	Number of Variables	Type of Variable & Potential Values
Frequency of Use of Assessments	<ol style="list-style-type: none"> <li>1. Times ELA is Used by Teachers of EB (Emergent Bilinguals)</li> <li>2. Times Running Records is Used by Teachers of EB</li> </ol>	2	<i>Ordinal Variable</i> , 0 times; 1-3 times; 4-6 times; 7-8 times; More than 8 times
Type and Extent of Knowledge Acquired	<ol style="list-style-type: none"> <li>1. ELA Pedagogical Content Knowledge</li> <li>2. ELA Student Knowledge</li> <li>3. ELA Assessment Knowledge</li> <li>4. ELA Bilingual Knowledge</li> <li>5. RR Pedagogical Content Knowledge</li> <li>6. RR Student Knowledge</li> <li>7. RR Assessment Knowledge</li> <li>8. RR Bilingual Knowledge</li> </ol>	8	<i>Ordinal Variable</i> , Not at all; A little; A moderate amount; A lot; A great deal
Frequency of instructional practice used as a result of ELA analysis of data	<ol style="list-style-type: none"> <li>1. ELA Use &amp; Grouping for Guided Reading</li> <li>2. ELA Use &amp; Prediction</li> <li>3. ELA Use &amp; Questioning</li> <li>4. ELA Use &amp; Summarizing</li> <li>5. ELA Use &amp; Analysis of Text</li> <li>6. ELA Use &amp; Sight Words</li> <li>7. ELA Use &amp; Decoding</li> <li>8. ELA Use &amp; Limited Vocabulary Texts</li> <li>9. ELA Use &amp; Teaching of Required Vocabulary</li> <li>10. ELA Use &amp; Figuring out Meanings of Unknown Words</li> <li>11. ELA Use &amp; Activating Prior Knowledge</li> <li>12. ELA Use &amp; Providing Students with New Background Knowledge</li> <li>13. ELA Use &amp; Providing Native Language Texts</li> <li>14. ELA Use &amp; Speaking about Texts</li> <li>15. ELA Use &amp; Partner Talk</li> <li>16. ELA Use &amp; Assessment Practice</li> <li>17. ELA Use &amp; Further Observation of Student</li> </ol>	17	<i>Ordinal Variable</i> , This assessment does not impact my use of this strategy; 1-3 times a month; 1-2 times a week; 3-4 times per week; Every day

**Continuation of Table 7. Overview of Variables used in Quantitative Analyses**

Type of Variable	Variables	Number of Variables	Type of Variable & Potential Values
Frequency of instructional practice used as a result of running record (RR) analysis of data	<ol style="list-style-type: none"> <li>1. RR Use &amp; Grouping for Guided Reading</li> <li>2. RR Use &amp; Prediction</li> <li>3. RR Use &amp; Questioning</li> <li>4. RR Use &amp; Summarizing</li> <li>5. RR Use &amp; Analysis of Text</li> <li>6. RR Use &amp; Sight Words</li> <li>7. RR Use &amp; Decoding</li> <li>8. RR Use &amp; Limited Vocabulary Texts</li> <li>9. RR Use &amp; Teaching of Required Vocabulary</li> <li>10. RR Use &amp; Figuring out Meanings of Unknown Words</li> <li>11. RR Use &amp; Activating Prior Knowledge</li> <li>12. RR Use &amp; Providing Students with New Background Knowledge</li> <li>13. RR Use &amp; Providing Native Language Texts</li> <li>14. RR Use &amp; Speaking about Texts</li> <li>15. RR Use &amp; Partner Talk</li> <li>16. RR Use &amp; Assessment Practice</li> <li>17. RR Use &amp; Further Observation of Student</li> </ol>	17	<i>Ordinal Variable</i> , This assessment does not impact my use of this strategy; 1-3 times a month; 1-2 times a week; 3-4 times per week; Every day
Demographic information	<ol style="list-style-type: none"> <li>1. Gender</li> <li>2. Type of Bilingual Setting</li> <li>3. Grade Taught</li> <li>4. Total Years Teaching</li> <li>5. Total Years Teaching Emergent Bilingual Students</li> <li>6. Certification Held</li> <li>7. Bilingual Extension</li> <li>8. Level of Education</li> </ol>	8	<i>Categorical Variables</i>  Potential Responses Vary Depending on Given Variable

The variables for instructional practice merit further discussion. They were developed in two ways. First, I consulted the literature to identify types of reading instructional practices that are considered most effective for emergent bilingual students (August & Hakuta, 1998; August & Shanahan, 2008). The types of reading instructional practices identified include:

- 1) teaching meta-cognitive strategies
- 2) adjusting the level of English vocabulary and structure
- 3) teaching word strategies
- 4) building background knowledge
- 5) oral language and
- 6) systematic assessment (August & Hakuta, 1998).

It is important to mention that these are broad categories of instructional practices. For example, the category of “teaching meta-cognitive strategies” includes strategies such as questioning, visualizing, and making connection between texts among others. In addition to these instructional categories, I included some instructional practices that came out my experience as a bilingual teacher in consultation with other experienced dual language bilingual teachers. These include:

- 1) regrouping students into guided reading groups
- 2) providing support in students’ home language.

From these broad instructional categories as defined by the literature and in consultation with experts in the field, the survey featured 17 distinct instructional practices that can be categorized into 8 instructional categories (see Table 8). For example, the category of meta-cognitive strategies is comprised of four instructional practices which include 1) prediction, 2) generating questions, 3) summarizing, and 4) analyzing texts. Table 8 outlines the instructional categories and each of the instructional practices featured in the survey.

**Table 8. Instructional Practices by Category**

<b>Instructional Practices</b>	
<b>1. Grouping</b>	
	Regroup students into guided <i>reading groups</i> based assessment data.
<b>2. Reading Strategies</b>	
	Ask students to make <i>predictions</i> prior to reading.
	Ask students to <i>generate their own questions</i> about texts.
	Ask students to <i>summarize</i> texts.
	Provide students with opportunities to <i>analyze</i> texts.
<b>3. Word Attack Skills</b>	
	Teach students <i>sight words</i> .
	Teach students <i>decoding strategies</i> .
<b>4. Vocabulary</b>	
	Provide students with texts with <i>limited vocabulary</i> .
	Teach students <i>required vocabulary</i> for reading.
	Teach students how to figure out the <i>meaning of unknown words</i> (ex, synonyms, context clues and cognates).
<b>5. Background Knowledge</b>	
	<i>Activate prior knowledge</i> before reading.
	Provide students with <i>new background knowledge</i> and experiences.
<b>6. Native Language Support</b>	
	Provide students with texts in their <i>native language</i> .
<b>7. Oral Language</b>	
	Ask students to <i>individually speak</i> about what they read.
	Engage students in <i>partner talk</i> about their reading.
<b>8. Assessment</b>	
	Provide students with more <i>opportunities to practice</i> this assessment.
	Further <i>observe and document</i> students' reading.

These instructional practices were worded in a manner that is familiar to teachers based on my experience of being a classroom teacher and a coach as well as on results of the cognitive interview. Secondly, the list of instructional practices was reviewed by two educational professionals with approximately ten years of experience each. This professional review helped to refine and confirm the wording of the instructional practices.

Lastly, it is important to mention that all of the variables that capture frequency of assessment use, teacher knowledge, and instructional practices were *ordinal* and used a likert scale. Ordinal variables represent an arbitrary ranking that is set forth by the wording of the survey. For example, any variable that gathers the extent or the perceived frequency of a given action is ordinal. In the case of this survey, all of the aforementioned variables quantify teachers' experience. They represent teachers' perceptions and memory of how often they use a given assessment, how much they perceive learning, and how often they implement a particular instructional practice. The wording of the measures either ask respondents to note frequency or level of knowledge gained. These types of measures were favored above measures of agreement-disagreement that are frequently used in surveys because these measures (agree-disagree) can only be analyzed in two categories (Fowler, 1995: 66). Rather, responses that measure frequency and level employ a numerical scale. The use of numerical scales yield fine-tuned information and makes the survey a more powerful instrument for analysis.

The demographic variables are categorical variables. In other words, these variables represent distinct groups such as either being female or male. These variables are free of ranking, order or numerical value. Before discussing the how each of the samples for the

qualitative and quantitative portions of the research were collected, a description of the assessments that were featured in the research will be offered.

### **Part 3: Assessments in Focus**

Although four very commonly administrated assessments were selected to be the focus of a number of the survey questions — the ELA, running records, reading conferences, and the NYSESLAT —the analysis focused on two of them: the English Language Arts Exam and Running Records. The interview questions were open ended and focused on the assessments that teachers brought to the discussion.

The assessments featured in the survey were the following: 1) New York State English Language Arts Exam (ELA); 2) New York State English as a Second Language Achievement Test (NYSESLAT), 3) formal running records and 4) reading conferences. These four assessments were chosen for this study because of their widespread use and potential familiarity with a wide range of teachers within New York City public schools. Furthermore, each of the assessments selected are regularly administered to most emergent bilinguals. Of these four different assessments, two are classified as summative or large-scale standardized assessments: ELA and NYSESLAT. Two are classroom-based assessments or formative: formal running records and reading conferences. The standardized tests are mandated by the New York State. The classroom-based assessments are also part of the New York City's standardized curriculum, Children First, as part of the balanced reading program. Therefore, all bilingual teachers who are part of the study had familiarity all of the assessments described here and employed them in their classrooms. A brief description of each of the assessments is provided to give an overview of

the purpose of a given assessment, how it is constructed and administered, as well as when emergent bilingual students take it.

### **English Language Arts Exam (ELA)**

The New York State English Language Arts (ELA) test is administered annually to students Grades 3-8.<sup>16</sup> The ELA is a literacy test that assesses reading, writing, and listening. Therefore, this test is a literacy content test and is intended to measure reading knowledge and skills of proficient English speakers. The mission of the test according the New York State Testing Program's (NYSTP) test manual is, "to evaluate the implementation of the State's learning standards at the student, school, district, and State levels (2010: 3)." This test is constructed through the collaboration between selected teachers and administrators from New York State as well as professionals from CTB-McGraw Hill. Final test questions are evaluated for alignment to state standards, appropriate contexts given student population and age, breadth and depth of material covered, as well as clear and concise language (NYSTP, 2010).

Students are administered the state exam in May and results are reported during the summer. The results from the NYS ELA are reported to teachers and parents in performance levels. These levels range from 4 to 1. Level 4 on the exam represents, "meeting standards with distinction." Level 3 represents, "meeting standards." Level 2 represents, "partially meeting standards." Lastly, level 1 represents, "not meeting standards." Students who perform at level 1 are mandated for summer school and are potentially candidates for retention. Emergent bilingual students who have been in the United States for less than a year on April 1 are exempt from taking the test. All other emergent bilinguals must take the exam. The following

---

<sup>16</sup> See Appendix 4 for an excerpt of the grade 4 ELA exam.

accommodations are available to emergent bilinguals: time extension, separate location, third reading<sup>17</sup> and bilingual dictionaries and glossaries (NYSTP: 10).

In addition, the goals for how ELA data should be used are stipulated in the Administrator’s Manual, the New York State English Language Arts (ELA) exam has three purposes: 1) “help schools identify students for whom they need to provide additional academic assistance,” 2) “help schools measure the effectiveness of their instructional programs,” and 3) “measure yearly student progress” (<http://www.p12.nysed.gov/apda/ei/ela-math-guide-11.pdf>). Results from this exam are used for the following purposes: student retention decisions, school quality reviews, and school progress reports.<sup>18</sup>

### **New York State English as a Second Language Achievement Test (NYSESLAT)**

The New York State English as a Second Language Achievement Test (NYSESLAT) is “designed to measure the English language proficiency of students who have been identified as English language learners” (NYSTP: 1). Students are identified as emergent bilinguals through the Language Assessment Battery – Revised (LAB-R), an assessment that is given to students upon entry to the public school system. The LAB-R determines whether a student is classified as an English Language Learner or not. If classified as an English Language Learner, the student is entitled to English language support. All emergent bilingual students are required to take the

---

<sup>15</sup> In the listening passage portion of the ELA exam, students hear the selected reading passage two times. Those students who are classified as English Language Learners are provided with the accommodation of having the passage read aloud three times.

<sup>18</sup> In New York City, school quality reviews (QRs) are two day visits from inspectors who evaluate the school according to pre-determined criteria. One of those dimensions is the school’s use of data. The QR focuses in on how data from both formative and summative assessments are analyzed and used by school administrators and teachers. School progress reports offer a snapshot of student performance and progress (gains or losses as determined by a comparison between results of the ELA from one year to another). Results are reported from A – F and made public on the Department of Education website.

NYSESLAT yearly to determine their growth in English language proficiency. The NYSESLAT is not a timed test and students have as much time as needed to complete it. Emergent bilingual students with special needs are administered accommodations as specified by their Individualized Education Plan (IEP).

According to the New York State Testing Program’s School Administrator’s Manual for the NYSESLAT, it “has undergone strict and extensive content and statistical analyses to ensure its validity and reliability (NYSTP: 1).” The test form is available in 5 grade bands – K-1, 2-4, 5-6, 7-8, 9-12 (NYSTP: 1). The NYSESLAT assesses students’ proficiency in four language modalities: listening, speaking, reading, and writing. Student performance on the test determines whether they remain classified as an “English Language Learner” or are re-classified as “English proficient.” NYSESLAT results are reported in proficiency levels. These levels are: beginning, intermediate, advanced or proficient. Student performance levels determine the amount of English language instruction support (in hours) that schools are mandated to provide students. Students are administered the NYSESLAT in May. Results are available to teachers in the Fall. Although this assessment was featured in the survey, the analysis of these results is not presented in this manuscript.

### **Running Records**

Formal running records are, “a systematic way of coding reading behavior so that the teacher can later analyze the behavior and make hypotheses about the use of strategies (Lyons & Pinnell, 2001: 38).” Through running records, teachers analyze students’ comprehension, fluency, miscues (reading errors), and use of reading strategies.<sup>19</sup> Miscue analysis is the core of

---

<sup>19</sup> See Appendix 5 for an example of controlled, leveled text and a recording sheet from a formal running record kit.

analyzing running records. Miscue analysis is the study of the type and percentage of errors that students make while reading. Some mistakes or miscues that students make interfere with the meaning of the text, while others may make sense but do not reflect the letters of the word printed in the text. Analyzing the type of errors made, helps teachers to target reading instruction. The miscue percentage is calculated by number of words correct divided by 100 words read. The miscue percentage determines if the text can be read by the student independently or is an instructional text (able to be read with the guidance of a teacher).

In addition to the analysis of student reading behavior and skills, teachers determine an appropriate independent reading level for students based on this assessment. Formal running records use controlled texts and recording forms that are part of an assessment kit. Controlled texts refer to texts exclusively used for the assessment and ones in which the level of difficulty is established by the publisher. Usually these controlled texts are shorter than published books students read. Each text is accompanied by a recording form that teachers fill out while a student is reading. This form aids teachers to determine reading strategies used by students and to calculate their miscue percentage and comprehension level. Both these calculations help teachers determine the difficulty level at which a student can read independently.

The Developmental Assessment Record 2 (DRA 2) is one example of a formal reading assessment kit. According to the publishers of the DRA 2, the kit ensures, “internal consistency reliability, passage equivalency, test-retest reliability, inter-rater and expert rater reliabilities” (Pearson: 2). Emergent bilingual students are administered formal running records. Each school may have a different policy of when emergent bilinguals’ reading levels in English are assessed through formal running records. Emergent bilingual students in dual language bilingual schools

are administered formal running records in both languages. Formalized running records are usually administered two or three times a year and the results, in reading levels, are reported to administrators and parents. These assessments are usually administered at the beginning, middle and end of the year. Results, in the form of reading levels, are available immediately. Other results such as those stemming from teachers' analysis of miscues are available to teachers, once they are done.

### **Reading Conferences**

Reading conferences are similar to formal running records. However, in a reading conference, teachers do not use controlled texts to assess students. Instead, reading conferences utilize the same texts that students have independently selected from a classroom or school library. Therefore the books that teachers use to conference with students may not be accurately leveled. Teachers also record student miscues and behaviors informally rather than through the use of a specialized recording form, as in the case of formal running records. A reading conference differs from a formal running record in that the teacher examines the student's day to day reading behaviors and also can inquire about the student's reading interests and disposition. There is a wide variation in the number of reading conferences teacher conduct. Emergent bilingual students can be conferenced with at various analysis levels of English proficiency to determine their reading levels, strategies used, language abilities, reading interests, and disposition. In dual language bilingual programs students would be conferenced with in their two languages. The results from these conferences are usually documented by teachers and used for future instruction. Conferences can occur during independent reading or while a student is reading a book that she or he has selected independently. Another possible time to conduct a

conference is during guided reading. Guided reading is when a teacher groups students who have similar instructional needs and provides them with the same text to read independently. This assessment although included in the survey and mentioned by teachers in the interview is not featured in the final analysis of this dissertation.

#### **Part 4: Sampling, Participants & Research Sites**

##### **Interviews**

More than 40 schools in four boroughs (Manhattan, Bronx, Queens and Brooklyn) were contacted in order to gain access to teachers for interviews. These schools were selected from a list of city-wide schools that have bilingual programs. Only those schools that had Spanish bilingual programs were contacted. This list of bilingual programs is readily available to the public from the NYCDOE's website. Out of the original number of schools contacted, five schools gave me permission to contact teachers to interview them. Interestingly, many of the 40 schools who were initially contacted stated that they only had transitional bilingual programs in the early grades (K-2), not in grades 3-5, and therefore could not participate in the study. Therefore, all schools where teachers were interviewed worked in dual language bilingual programs, which in the end were the only schools that had any type of bilingual program in grades 3-5. It is important to note, that dual language bilingual schools are often referred to "model programs" because they advance both bilingualism and biliteracy. In theory, if any teachers are using assessment to its full potential for emergent bilingual students, it would be teachers in these types of programs.

After receiving access to a school through the principal's permission to conduct research in the building, flyers describing the research and soliciting participants were provided to either

the assistant principal or coaches. These recruitment flyers were then physically given to teachers and emailed to them. In one school, the principal provided one “prep” or 45 minutes of non-instructional time to teachers so that they could hear my presentation of research goals before being recruited.

Teachers from the schools either contacted me or the liaison at the school to set up a meeting time for the interview. Although five school principals gave me permission to conduct research at the school, interviews were conducted at four schools. At one of the schools, no teachers volunteered to be interviewed.

Interviews were conducted at four schools. All teachers interviewed taught in Spanish-English dual language bilingual programs. Two of the schools were entirely dual language bilingual schools, while the other two were schools that housed these programs within them.<sup>20</sup>

Table 9 describes the demographic and socioeconomic characteristics of the schools. A narrative description of each of the school’s bilingual programs follows.

**Table 9. Demographics & Socio-economic Characteristics of Schools**

School	Total Student Population	Grade Levels	% Free School Lunch	% English Language Learners	Percent Hispanic	Percent Black	Percent White	Percent Asian
#1*	689	K-5	67	14	51	26	14	5
#2	436	preK-8	96.1	44.04	97.5	2.5	N/A	N/A
#3	420	K-8	74.6	26.9	85	1.8	8.5	3
#4*	1161	K-4	87.5	50.2	83.6	1.6	4.6	9.7

<sup>20</sup> Schools with dual language bilingual programs within them are denoted by the asterisk in Table 6.

## **School 1**

This school is located in a neighborhood that draws in high, middle, and low income students into its total population. The dual language bilingual program is balanced, meaning that about half of its students' home language is English and for half it is Spanish. This dual language bilingual program has both general education and Integrated Co-Teaching (ICT) classes.<sup>21</sup> The bilingual general education classroom has one head teacher, while the bilingual ICT classes have two head teachers: one general education and one special needs teacher. There are two classrooms per grade. The program is a “one-day-one day” model meaning that the language of instruction alternates every other day.

As this study focuses on assessment, the assessment reports about each school generated by the NYCDOE will be featured for each of the schools. School Progress Reports, a document summarizing each school's testing performance for a given year, provide a snapshot of the school's testing success to the public. They offer information about student performance – or the percentage of students who attained passing grades on the New York State tests for a given year – and student progress – or the percentage of students who made one year's worth of gains according to the state tests. According to the school's 2010-2011 Progress Report issued by the NYCDOE, School #1 earned “Cs” on both student progress and performance. The grade for school environment was an “A.” It is important to remember when looking at these measures, that the grades that schools receive represent both English literacy and math exams.

---

<sup>21</sup> Integrated Co-Teaching, commonly referred to by the acronym ICT, is a classroom where general education students learn alongside special needs students. These classrooms have two head teachers – one certified in general education and the other in special education. In bilingual ICT classes, both teachers are also certified in bilingual education. This inclusion model was formally referred to as Collaborative Team Teaching (CTT).

Furthermore, these grades do not disaggregate students in the dual language bilingual program versus those that are not. The school was given an overall grade of “C.”

I met with six teachers individually from this school. They were interviewed either after-school or during lunch periods.

## **School 2**

School 2 is located in a predominantly Latino neighborhood. Many of the students’ families have emigrated from the Dominican Republic. However, this school attracts students from a neighborhood which has mostly middle to high income families whose home language is predominantly English. As a small school, there are only two classrooms per grade. There is one head teacher for each classroom.

The language allocation policy for this school differs depending on the grade level taught. In the early grades, Kindergarten and first grade students are instructed in Spanish on Mondays, Wednesdays and Fridays. They are taught in English on Tuesdays and Thursdays. Starting in the second grade and continuing to the fifth grade, an alternating half-day cycle is implemented. Students, for example, are taught in Spanish one morning, in the afternoon they are taught in English for a period of a week. The following week, they are taught in English in the morning and then switch to Spanish in the afternoon. The length of the week cycle varies according to the grade. In middle school, language arts, math, science and social studies are taught in two week language cycles: two weeks in English and two weeks in Spanish.

This school received an overall grade of “A” on the school’s Progress Report for school year 2010-11. Additionally, the school received “Bs” for both student progress and student performance. Lastly, the school was issued an “A” for school environment.

Four teachers were interviewed individually at this school. They elected to be interviewed either during their free period or lunch.

### **School 3**

School 3 is located in a brand new building in a predominantly Latino neighborhood. Most of the Latino population hails from the Dominican Republic and the Caribbean. Many students who come into the program have Spanish as their home language, although there are a small percentage of students who have English as their home language. The school also has started a bilingual ICT program and therefore has both one bilingual general education and ICT class on each grade.

The dual language bilingual program's language allocation is as follows: on Mondays, Wednesdays and Fridays all instruction is in Spanish; on Tuesdays and Thursdays instruction is conducted in English. Special classes such as art, music, and gym are in English. Therefore, although students receive three days of instruction in Spanish, because of the special classes are in English even during those days, school leaders believe that the time allotted for each language balances out to fifty percent for each.

This school received an overall grade of "C" on the Progress Report for the 2010-11 academic year. Student performance was allotted a grade of "C" while, student progress was deemed a "D." The school was awarded an "A" for measure of school environment.

Five teachers were interviewed at this school. Interviews were conducted in focus groups after school.

#### **School 4**

School 4 has the largest school population out of the all of the schools that participated in the study. It is also located in a predominantly Latino neighborhood where many students come from Mexico and South American countries. The dual language bilingual program is housed within the larger school. This school also is beginning to attract middle to high income families whose home language is predominantly English into the student population. In addition, teachers reported that a few students whose home language is Chinese are enrolled in the two-way English-Spanish immersion program. Despite this, the majority of students come into the program only knowing Spanish.

The dual language bilingual program was also comprised of two classes per grade. However, at this school there was one teacher who was designated to teach in English and one teacher designated to teach in Spanish. Students alternate between the Spanish class and the English class every other day.

According the school's Progress Report for 2010-11, the school received an overall grade of "A." Student progress and performance was deemed a "B" for both measures. Finally the school received an "A" for environment.

At this school, interviews were conducted during preps provided to each grade level during the school day. Teachers had the choice whether to use the prep or not. Therefore, some of the interviews at this school were conducted in pairs with the English and Spanish teachers for each grade present. A total of five teachers were interviewed at this school for the study.

In sum, the four focal schools from which teachers were interviewed demonstrate differences in size, population, and language allocation policy. Rather than seeking similarity,

diversity in setting was intended for this piece of the study. It is this difference between dual language bilingual schools which can provide the foundation for analysis of themes across schools, despite the particular contextual factors across the school sites.

Teachers were recruited solely by flyer and interest. After more than three-fourths of the interviews were conducted, a small grant was acquired. A portion of these funds were used to retroactively provide teachers interviewed with a small stipend. This retroactive stipend matched the incentive offered (\$10 gift card to Barnes and Nobel's) to teachers who were recruited to be interviewed after the grant was obtained. Therefore, in both the recruitment of schools to participate in the study and in the teachers who were interviewed, convenience sampling was used.

Those teachers who were interviewed were eager to speak about assessment of their emergent bilingual students and spoke with candor about their challenges in assessing students. Their commitment to the project was evidenced by the time they spent during their free periods – planning periods, and lunch – in order to talk to me about the issues.

A total of 20 teachers out of the original goal of 24 teachers were interviewed. These teachers ranged from having one year to thirteen years of experience. Teachers interviewed averaged about seven years of experience teaching. Half of those teachers interviewed were fourth grade teachers. Fifth grade teachers were interviewed the least. Furthermore, 65% of teachers interviewed held permanent certification. Table 10 displays demographic information about the teachers from the interviews. This information is displayed alongside demographic information about the survey sample to offer a comparison between the two samples.

**Table 10. Descriptive Statistics for Sample of Teachers Interviewed and Teachers Surveyed**

	<b>Teachers Interviewed</b>	<b>Teachers Surveyed</b>
<b>Sample Size</b>	20	63
<b>Total Years Teaching (Mean and Standard Deviation)</b>	6.6 (M) 3.7 (SD)	9.8 (M) 8.8 (SD)
<b>Years Teaching Bilingual Students (Mean and Standard Deviation)</b>	6.3 (M) 3.1 (SD)	8.3 (M) 6.9 (SD)
<b>Grades Taught – Percent of Sample</b>	3 <sup>rd</sup> Grade – 35%  4 <sup>th</sup> Grade – 50%  5 <sup>th</sup> Grade – 15%	ESL across grades– 11.1%  3 <sup>rd</sup> Grade – 31.7%  4 <sup>th</sup> Grade – 20.6%  5 <sup>th</sup> Grade – 15.9%
<b>Type of Program for emergent bilinguals – Percent of Sample</b>	Dual Language Bilingual– 100%  Transitional Bilingual – 0%  English as Second Language – 0%	Dual Language Bilingual – 64.15%  Transitional Bilingual – 22.64%  English as Second Language – 13.21%
<b>Type of Certification – Percent of Sample</b>	Permanent – 65%  Alternative – 10%  Initial – 25%  None – 0%	Permanent – 56.36%  Alternative – 9.09%  Initial – 32.73%  None – 1.87%
<b>Bilingual Extension – Percent of Sample</b>	Yes – 87.5%  No – 12.5%	Yes – 85.45%  No – 14.55%
<b>Educational Level – Percent of Sample</b>	BA – 25%  MA – 75%	BA – 7.27%  MA – 92.73%

## Survey

For the quantitative work, random sampling was impossible. Random sampling would involve randomly selecting a sample of the total population of bilingual teachers in New York City. This type of sampling would allow the results of the survey to be generalized to the entire population of New York City bilingual teachers. However, the challenge was in acquiring the database of all the bilingual teachers from the New York City Department of Education which is not available to researchers. Instead, snowball sampling was used. Snowball sampling is a method by which a group of people known to the researcher is contacted. This group then sends the survey to their list of contacts in hopes of finding respondents who fit the criteria required to fill out the survey.

Surveys were administered in two ways. Surveys were available to participants at the Dual language bilingual Symposium sponsored by the New York City Department of Education in April 2011. Fifteen surveys were collected in this way. In addition, the survey was available on Survey Monkey. The link to the survey was distributed three times between April-June 2011. The survey link was sent to a wide range of educators – school administrators, teachers, professors and bilingual education activists – to distribute to their contacts. Surveys were filled out by dual language bilingual, transitional bilingual and ESL teachers.<sup>22</sup> Furthermore, the acquisition of a small grant allowed me to offer survey respondents a small incentive (a \$5 Staples Gift Card was offered).

---

<sup>22</sup> For a list of the types of teachers who work with emergent bilinguals and which were included in the survey and the interview, see Appendix 6.

In order to calculate the response rate, contacts who were sent the survey link gave an estimate of how many people they sent the link to. Forty-eight surveys were collected through survey monkey, which resulted in a response rate of 18.9%.<sup>23</sup> Unfortunately since the response rate was so low, the data could not be used to generalize the population of 3-5 grade bilingual teachers. Assessment as a topic is a sensitive one to teachers. Although teachers are often critical of both summative assessments and formative assessments for different reasons, the use of assessment within schools is often unquestioned in the public realm. Therefore, perhaps the low response rate is due to the sensitivity which surrounds this topic.

Table 10 displays the demographic information about the sample surveyed, alongside the teachers interviewed. The sample of teachers surveyed averaged about 10 years of teaching experience teaching, about 3 more years on average than those teachers who were interviewed. In addition, more than 60% of the teachers surveyed taught in dual language bilingual programs. About 20% of teachers taught in transitional bilingual programs and about 13% in ESL programs. In contrast, all teachers interviewed taught in dual language bilingual programs. Both of these samples do not reflect the number of dual language bilingual programs that exist in the city. As described in the introduction, only 4% of emergent bilingual students who are designated as English Language Learners by the NYCDOE are in dual language bilingual programs.

---

<sup>23</sup> Response rate was calculated by estimating the number of people who received the Survey Monkey link via email. It was estimated that approximately 254 teachers received the link to the survey.

## **Part 5: Data Analysis**

Data analysis began in July 2011. For the analysis, qualitative and quantitative data was examined separately at first. In order to interpret the findings, the results from these two portions of the study were analyzed separately and then in combination with each other. It is through this weaving of methods during data analysis, that a full and detailed view of the processes by which teachers generate knowledge and translate them into instructional practices for emergent bilingual learners was constructed. Although the intent was to weigh the qualitative and quantitative data with equal measure in the analysis, because of the response rate, the main findings of the study incorporated the quantitative data analysis, but not on equal standing as the interview data.

### **Interviews**

All respondents gave permission to digitally record interviews. All recordings were transcribed. Prior to analysis, codes that paralleled some of the quantitative variables were identified. These were codes for types of knowledge (content, student, assessment and bilingual) and instructional practices. However, the majority of final codes that were used in the analysis emerged from the data (see Appendix 7).

The entire body of transcribed interviews was read several times to gain a sense of topics that were discussed. Subsequent readings had specific goals for data analysis. One reading was intended to pull themes from the interviews. These themes emerged from the interviews rather than the research questions. One such theme, for example, was teachers' use of assessment data bilingually. Within this theme teachers talked about language transfer and the inequities between

assessment materials in English and Spanish. Topics within these themes were later translated into codes.

Further readings were done in order to develop codes. Codes stand for discrete ideas and concepts that were found within the teacher interviews. For example, teachers may analyze a running record with respect to oral language, punctuation, and fluency. Each of these ways of analyzing a running record was given a code. In addition, the frequency that codes appeared throughout the interviews was noted, to provide information about how often each of the themes emerged for this group of teachers.

Lastly, the interviews were scanned to identify examples and descriptions of how teachers develop knowledge from assessments for emergent bilingual students. These key quotes were highlighted and used in the final analysis.

## **Survey**

Data from the survey was reviewed and organized prior to being uploaded to both STATA and SPSS, statistical software packages used to analyze the data. Both STATA and SPSS were used in the analysis in order to benefit from the advantages each offers in terms of programming.

First, the data was simply analyzed for trends – what types of knowledge teachers report they gain from working with assessments of emergent bilinguals and what types of instructional practices teachers most frequently engage in after consulting the assessment data of emergent bilinguals. Assessment use was analyzed in order to gain a sense of assessment practices that teachers surveyed have.

Subsequently, analyses were conducted to answer the research questions posed in this thesis. In order to answer questions — What kinds of knowledge do teachers gain from examination of reading assessment data? and What changes in reading instructional practices result from teachers' use of assessments? — percentages of two categories of variables were calculated. As noted previously, variables for these categories – types of knowledge acquired from assessment of emergent bilinguals and instructional practices implemented after examining assessment data – are assessment specific. That is, one set of variables refer to types of knowledge and types of instructional practices that are a result of the ELA. The other set of variables refer to types of knowledge and types of instructional practices acquired from running records. This allowed for comparisons for each variable between these two types of assessments. Cumulative distribution graphs, to be described further in the findings, were constructed to display differences in responses for each variable between the ELA and running records. In addition, frequency tables also describe each of the variables. Lastly, a non-parametric test, the Wilcoxon signed rank test was used to determine if statistical differences existed between the variables for teacher knowledge and instructional practices resulting either from ELA or running record analysis.

In order to determine how teacher knowledge was related to assessment use and how instructional practices were related to teacher knowledge, correlation analysis was employed. Correlations are defined as the degree of association between variables (Glass & Hopkins, 1996: 103). Correlations can fall between a value of -1 and 1. The value of the correlation signals the strength of the association, while the sign of correlation represents the direction. For example, a correlation of .8 would represent a strong positive correlation between variables.

Although correlation describes the relationship between variables, it cannot explain the relationship the two variables causally (Glass & Hopkins, 1996). However, correlation coefficients once placed alongside qualitative data, may be able to point to reasons or causes why a particular relationship may exist and its strength. Although the combination of the qualitative and quantitative data may begin to point to reasons for causality, this research is not designed to answer causal questions.

## **Part 6: Limitations**

### **Interviews**

As a teacher and coach who worked in a dual language bilingual education program for the past 12 years, I am very invested in this work. This experience was an advantage in establishing relationships with teachers that I interviewed. In fact, this experience possibly lowered the possibility of reactivity, or the effect of my presence on the research, since I occupy both the position as a teacher and an outside researcher, not an official from the Department of Education (Maxwell: 109). However, my history as a bilingual teacher may introduce bias into the research. I hold strong assumptions that bilingual education is an important option for bilingual students. Although I hold these views, I also have a commitment to making sure that the research “speaks for itself.” Although my viewpoint comes out in any rendering of the project, I hope to make clear to readers when my opinion is at the forefront and also to present data in such a way to reveal itself to readers.

One limitation with the interview data is respondent selection. Due to the difficulty of gaining access into schools, schools were selected, or rather allowed me to enter their school building based on a principal who was open to having research done on assessments in their

building. At this moment in time, when assessments determine the quality of a school, the leadership and teachers, this can be a sensitive topic. Furthermore, teachers who volunteered to be part of the research for the most part did so without an incentive (incentives were provided to a few participants, others received a retroactive small “stipend”). Therefore, perhaps, the sample of teachers who decided to be interviewed was particularly motivated by their feelings about assessment of emergent bilinguals in order to volunteer to be part of this project.

Furthermore, all of the teachers interviewed were dual language bilingual teachers. Therefore, they do not represent teachers of emergent bilinguals across the city. The analysis from interviews can possibly only be generalized to the small sample of dual language bilingual programs in the city.

## **Survey**

One of the factors affecting the validity of the survey is response rate and sampling (Judd, Smith & Kidler, 1991, 216). Response rate refers to the percentage of surveys completed. A response rate should be sufficiently high to warrant comparisons to a given population. As previously noted, the response rate for this survey was 18.9%, thus eliminating the possibility that this data could be generalized to population of grade 3-5 bilingual teachers.

Additionally, the collection of 100 completed surveys was the goal set forth at the proposal stage. The final count of surveys collected was 63. Out of the 63 surveys gathered, eight were missing large amounts of data, thereby leaving 55 surveys that were “complete.” Furthermore, data was collected through snowball sampling. Like the interview data above, the survey data is open to bias based on respondent selection. Teachers who received the email were those who filled it out. Perhaps, the low response rate was due to the sensitive nature of the

topic. Likewise, the survey sample does not reflect the percentages of emergent bilingual students who are in different types of bilingual programs in the city. Rather, this sample has a high percentage of teachers who work in dual language bilingual programs.

Another limitation is the inability to tease out whether the instructional practices that teachers report using are ones that they would have used regardless of the use or knowledge generated from assessments. Although I am able to inquire about prior use of instructional practices through questioning in the interviews, I do not have an interview question in the survey that can adequately capture whether instructional practices are the result of looking at assessment data or are a regular part of teachers' instructional practices. Additionally, instructional practices may be affected by a range of other factors which are not accounted for by the survey. For example, the school administration may mandate certain repertoire of instructional practices. Although this information may come out through the interviews, no information about this factor or others is captured by the survey.

Furthermore, the survey data does not allow for different responses based on varying levels of proficiencies of emergent bilingual students. The survey asks teachers to respond in general when considering the assessments of their emergent bilingual population. However, emergent bilingual students are a diverse group, even within one classroom, and the variety of needs within this group may require diverse instructional practices for different sub-populations. Therefore, when teachers respond to questions about their emergent bilingual students, they may reference one sub-set of this group, i.e. their least proficient group of emergent bilinguals, and answer questions about instructional practices with this group in mind.

The sample size of the quantitative data collected therefore is small and poses limitations on both which analyses could be conducted and the extent to which these analyses can be interpreted. Despite these challenges, data analyses provide a picture into how assessments impact teacher knowledge and hence, may impact instruction. However, I am aware of the limitations in the analysis and I hope the analysis respects the limitations of the data.

### **Summary**

Mixed-methods were employed to answer the research questions posed for this doctoral research. These methods are best suited to this inquiry because through their use, both particularities and trends of how teachers learn from and use assessments of emergent bilingual students can be brought forth. Additionally, the presence of an overlap between the methods strengthens the study as it is through this commonality in the data that provides the ability to compare results from both types of analysis. The unique characteristics of the qualitative data collection allows for teachers to report both school-specific and personal anecdotes thus affording a personal perspective of the phenomena being studied. Although due to small sample size and low response rate the interpretation of the quantitative data is limited, those analyses will be presented to bolster qualitative findings.

## **CHAPTER 4:**

### **Results from Interviews**

In this section findings from interviews are presented. The goal of the qualitative analysis is to provide an overview of themes that emerged from the interviews with respect to the development of teacher knowledge and the practices that teachers report implementing after analyzing assessment data of emergent bilingual students. Another goal of this analysis is to offer readers an opportunity to connect with the voices of teachers in order to understand their perspectives on the role of assessment in their knowledge construction and in shaping instruction. The discussion is organized around the themes that emerged from the data. In this way, themes from this analysis go beyond the following research questions that were investigated through this data:

- 1) What kinds of knowledge do teachers of emergent bilinguals gain from the examination of reading assessment data of emergent bilinguals?
- 2) What changes, if any, in reading instructional practices result from the types of knowledge gained by teachers of emergent bilinguals from large-scale standardized assessments and classroom-based assessments?

In this chapter, I argue that the summative and formative assessments studied in this thesis do not fulfill the potential of providing teachers with a solid knowledge base of their emergent bilingual students' reading development. Although these two types of assessments are used and experienced differently by teachers of emergent bilinguals, with teachers being more critical of summative assessments than of formative assessments, they both fail to take into account student individual characteristics, especially their emergent bilingualism. Therefore, they both miss the mark in providing teachers with knowledge that is specific to the reading

development of emergent bilingual students. Rather, I contend that the use and analysis of both of these assessments by teachers is characterized by “ritualized assessment practices” or a fixed repertoire of behaviors that move teachers through a hurried process of collecting and “looking” at data without serious consideration of how student language development deeply impacts both the interpretation of assessment results and any instructional action plan that may result. Lastly, in both the case of the summative and formative assessment, instructional practices that are implemented are weakly the result of the knowledge teachers have acquired from the analysis of assessment data, if at all. Instead, instructional practices implemented reflect a set of *ritualized* behaviors that are *routinely* implemented by teachers and which are sponsored by schools.

This chapter is divided into two sections in order to provide a close look into the particular issues of how teachers use, learn, and experience both summative and formative reading assessments of emergent bilinguals. All of the quotes provided in this chapter identify the teacher through a pseudonym. Furthermore, the school where teachers work, the grade they teach, and the date that the interview was conducted are provided alongside the pseudonym. All interviews were conducted with teachers in dual language bilingual programs in New York City.

### **Part 1: Teachers Speak about the English Language Arts Exam (ELA), a Summative Assessment**

Students, grades 3-5, are given the English Language Arts (ELA) exam yearly. This summative assessment evaluates students’ literacy abilities through a variety of tasks. Students are presented with reading passages in different genres and asked to answer multiple-choice questions about these texts, write short responses to questions, fill in graphic organizers, and write essays. In addition, students are read aloud a passage by the teacher. Students are then asked to answer multiple choice questions and write an essay based on the listening passage.

Emergent bilingual students who have been in the country for one year or more must take the ELA. These students are often provided with accommodations such as extended time and bilingual dictionaries or glossaries. These types of tests, as discussed in the literature review, were created to assess the knowledge of students in English, and therefore may inaccurately measure the reading skills of emergent bilingual students whose home language is not English, even when these students are provided with accommodations. It has been argued that these high-stakes tests are inappropriate for use with this student population, producing results with questionable meaning (Valdez & Figueroa, 1994; Peterson & Rivera, 1994; Abedi, 2008).

Within the framework of test-based accountability, summative assessment data are supposed to direct instructional decision-making (Linn, 2003). As such, teachers are expected to analyze and use the results of the ELA. The ELA is administered in May and results are given to teachers at the beginning of the next school year. Fourth and fifth grade teachers review their incoming students' results.<sup>24</sup> Additionally, teachers, grades 3-5, may also review their outgoing students' results in order to assess how their students from the previous year fared.

Results are reported two ways. First, teachers receive each student's *performance level*, which is reported as a score from one through four, along with raw scores. While scores two and above are considered passing, scores of three and four are considered at and above grade level. Performance levels are the overall score that represents how students fared on all parts of the exam. Therefore, it is difficult to tease out from this one score how students did on reading, writing, and listening portions of the test. Furthermore, teachers receive a report of *performance*

---

<sup>24</sup> Third grade teachers whose incoming students were in second grade the year before do not have standardized test data to analyze.

*indicators*. Each performance indicator represents a cluster of questions included in the test that reflect a given reading skill. These skills include: finding the main idea, using supporting details, identifying author's purpose, and inferencing, among others. The report organized by performance indicators displays how students did on these types of questions on the test. The performance indicator report does not distinguish between reading, writing and listening — it solely lists students' performance by reading skill.

The relationship between ELA results of emergent bilinguals and teacher knowledge is both complicated and conflicted. As my analysis demonstrates, teachers believe that they should learn from the assessment results of the ELA, but when probed this is not the case. In fact, teachers express that they learn very little about their emergent bilinguals from the exam. The reason for teachers' lack of learning can be explained by their keen awareness of the limitations of this type of assessment to reveal the reading abilities of emergent bilinguals. This lack of trust in the assessment to yield valuable information affects teachers' readiness to acquire knowledge from the use and analysis of assessment data. Despite reporting that they do not acquire knowledge from ELA results of emergent bilinguals, teachers actively use the results to identify patterns of performance across students and sort them into instructional groups. These three topics — 1) teachers' critiques of this assessment, 2) the knowledge teachers acquire, or do not acquire, from ELA data of emergent bilinguals, and 3) the ways in which ELA results are used by teachers for instruction — will be discussed in the this section of the chapter. I end the section by providing teachers' accounts of how to rescue the ELA assessment as a way to learn about emergent bilingual students. I now turn to the teachers' critiques of using the ELA data to learn about emergent bilingual students.

## **A Lack of Trust: Critiques of the Use of ELA Data**

In order to understand why teachers are not learning about their emergent bilinguals from the ELA, it is important to examine their critiques, as these are the foundation for how ELA assessment data is ultimately used by teachers. Teachers' critiques ranged from the identification of confounding factors in the exam to the way in which tests are used to measure student progress from grade to grade. Lodged within these critiques is evidence of teachers' existing knowledge base of emergent bilingual students, knowledge developed through classroom interaction and not assessment per se.

A common critique of the ELA mentioned in the interviews is that it measures listening, writing, and reading simultaneously. For this reason, it is tricky for teachers to untangle the combination of factors which impact emergent bilingual student performance on the test. For example, teachers describe that students listen to a passage and subsequently answer multiple choice questions and write at length about that passage. Emergent bilingual students who are capable of understanding a listening passage, but whose reading and writing abilities are not at the same level as their listening skills, may not be able to demonstrate their ability to extract information from a listening passage. Therefore, teachers noted that it is difficult for them to tease out if a student's final score on a large-scale standardized exam is attributed to their skills in reading, writing, listening or a combination of the three. With regard to the confounding of different literacy abilities, this 4<sup>th</sup> grade dual language bilingual grade teacher states:

We don't know ... [with the ELA] you don't fully understand if this is a comprehension issue or for example, on the 4th grade test, there's so many short response and extended response questions, or if this is a writing issue. I don't feel that the ELA is informative. Especially for my English Language Learners, it's particularly difficult in the sense that I can usually attribute it to, or at least in my past experiences my English Language

Learners have suffered with their writing because there are so many different components. So the grammar is going to affect them, the spelling is going to affect them. On top of learning what an essay format looks like, they have to adapt to all these other things, the grammar component even to the vocabulary, so I assume that it's writing, but I wish with the ELA when we get their scores, that they are broken down more than just the raw score. The raw score doesn't really help me in that regard. (Amy, School #1, Grade 4, 11/26/10)

First, this teacher's quote highlights that many teachers do not believe that the test results are disaggregated in a meaningful way in order to extract information about students, regardless if students are emergent bilinguals or not. Specifically for emergent bilingual students, because the ELA "clumps" together a range of literacy skills, it is difficult for teachers of emergent bilingual students to attribute their performance to one literacy modality (listening, reading or writing) over another. The ways in which ELA data is reported to teachers is problematic because the test is not designed to offer student-level information.

Furthermore, this teacher provides important information regarding how bilingual teachers think about working with their language learners. When examining a writing piece derived from a reading assignment, teachers of emergent bilingual students must consider a range of factors, such as grammar, spelling, structure, and language use, among others. But the ELA does not give teachers the exact information about the students' writing that they consider essential. Rather, teachers employ both their knowledge of students derived from their personal classroom interaction, as well as their knowledge of language acquisition and development when forming their opinions about why the ELA does not accurately assess their emergent bilingual students.

This concern about the precision of this assessment to isolate and evaluate the reading abilities of emergent bilingual students was also voiced with relation to how teachers believe test items are linked to skills being tested. Teachers noticed that what is measured by the test is not clear-cut, especially for emergent bilinguals. Susanna, a former 5<sup>th</sup> grade teacher and currently a coach, says that test questions require a composite of literacy skills in order to be answered. Therefore, although the performance indicators are believed to capture how students fare on specific reading skills, performance indicators are far from being distinct. She states:

We noticed that our children struggled on that [a particular performance indicator] so we looked at what the question was specifically and we got the actual test booklet and looked at it. In many cases you find that there are so many more things going on in the question than the performance indicator tells you. So we are getting this data and we are saying we really stink at main idea and then you look at the question you can tell maybe they had to read a chart or maybe it was about synthesizing some other information and there were other things that went into that. So you can't really separate all these skills and test them individually, they go together in the thinking process, but in some ways we found that misleading that the test as a tool has some limitations because you really can't separate those processes in real life. (Susanna, School #2, Coach, 11/26/10)

This concern was repeatedly echoed by teachers. Suzie, a fourth grade teacher stated that test questions did not match the skill that they were supposed to be evaluating. Suzie says:

The ELA data I don't find particularly helpful. The categories of information that they break the ELA into seem kind of arbitrary in a way and one of the categories, I think it is inferencing or something. They'll say that these questions tend to be about inferencing but then if you look at the questions themselves, they don't always seem to match up with the categories the ELA test makers are giving to that question, so how is this data useful to me if I don't even trust the way that they are categorizing the data? (Suzie, School #2, Grade 4, 11/12/10)

Both Susanna's and Suzie's narratives demonstrate that when teachers engage in inquiry about how the test construction matches its purpose, the test falls short. In doing this, they are building

knowledge of the limits of this assessment. Their quotes raise serious concerns regarding the way that test results are reported. If the test results are not disaggregated in valuable ways to teachers, then teachers will not consult these reports to make sense of their students' performance.

As suggested by both Susanna and Suzie, in order for the ELA exam to be beneficial to teachers of emergent bilinguals, the exam must be examined in detail in order to understand how test items evaluate student learning. Above all, it is important for teachers to consider the language demands that test items may present to emergent bilingual students and the impact that language may have on how students answer test questions. Although teachers did not engage in this type of analysis of the ELA routinely, it is important that teachers build assessment knowledge of large-scale exams and hold it alongside student knowledge, as well as knowledge of language acquisition and development so that their analysis of the test is mindful of students' language abilities and how these abilities affect their performance.

Teachers also have serious concerns about how emergent bilinguals at different stages of English language development may fare on the test. Many of the teachers interviewed remarked that especially for students who are recent arrivals and who are mandated to take the annual large-scale standardized tests after just one year in the US, the results are troublesome. These students do not have language proficiency to truly demonstrate their literacy abilities on the test.

Emeline, a third grade teacher summarized it succinctly this way, "I think the ELA is for ELLs who passed the NYSESLAT. Because otherwise, [the results show] this is how much you don't know in English because they are being compared to a bunch of other kids who are proficient" (School #1, Grade 3, 11/7/10). This argument was repeated by many of the teachers

interviewed and cited as one of the key reasons why ELA test results were not considered appropriate for these students by teachers.

Susanna confirms Emeline's assertion that assessment data for recent arrivals is often not taken into consideration by teachers. She provides a rationale for the purpose of assessments in general — to reveal the limits of student knowledge thus demonstrating what they can do rather than focus on what they don't know. Hence, Susanna's comment below also offers a call for further informal assessments:

Especially when we get new students and they start testing after a year, I feel like we always just disregard it, we know the student is not at grade level in English yet. We want assessments so we can find out where they are and this is not going to help us. It's just going to tell us that they are not there. So what we want to know is where they are and what their next steps are and so that's when informal assessments are more important. (Susanna, School #3, Coach, 11/26/10)

Teachers also noted that these assessment results are murky for emergent bilingual students because these students have few test-taking skills given that they may be new to testing. This lack of test-taking skills may play a large part in emergent bilingual students' final scores on standardized assessments. The following teacher, Aura, at School #1 affirms:

You know, it's tricky because it's the answer that they have to bubble in, if a student doesn't know how to do that, then that's another issue. You know, so a lot of things can go on in the ELA that I worry about with my students. (Aura, School #2, Grade 3, 11/12/10)

This point is important because it highlights that test-taking behaviors that are not necessarily tied to reading, but rather are learned through practice and exposure, are factored in when students are assessed through large-scale standardized exams. This lack of experience and practice, although affecting many types of students, may weigh more heavily for emergent

bilingual students since test-taking behaviors may be thought of as “cultural.” Therefore, in addition to the confounding of different literacy abilities such as writing, reading, and listening on large-scale standardized tests, test-taking behaviors may also mask emergent bilingual students’ actual reading competencies, especially in the elementary grades when these skills are new to students.

Moreover, teachers were concerned about the types of passages included on the test. They noticed that passages featured were “foreign” to emergent bilinguals. This characteristic of test passages has important consequences for emergent bilinguals. When students are unable to connect or activate prior knowledge in their reading of text, the difficulty of that passage increases, and therefore comprehension may be compromised (Cooper et al., 2010). In essence, the same passage may be more or less difficult for a student with the same reading ability, given the level of background knowledge the student possesses about the topic. A fourth grade teacher at School #4 remarks:

The content of these passages that we assess them with, sometimes they are just unrealistic for the kids. They cannot apply any prior knowledge to it. There was a passage about being in Alaska and a frozen swamp and you know it’s vocabulary that they are not aware of and they can’t use their prior knowledge. Did the kid really not pass, or is it that the text is not the right text? So it becomes a gray area. (Marisol, School #4, Grade 4, 12/1/10)

Teachers’ critiques do not end at how the test is constructed. Teachers also commented that the pressure of doing well on the assessment lead to test-prep sessions. Many teachers attest that they spend a great deal of time assisting students to do well on these high-stakes standardized tests. A fifth grade teacher at School #2, who teaches in an Integrated Co-Teaching

(ICT) class, where special needs' emergent bilingual students learn alongside general education emergent bilingual students, is remorseful about the time spent teaching "test-prep." She says:

I felt that so much of our test prep was just a waste of time because it didn't ... these kids still failed and I wish that we had spent so much more time on rich, enriching activities, reading books, actual books with pictures and diagrams and stories. I wish we had spent more time doing that because, yes, test prep skills are important skills to have, but for these kids they are not there yet and it's just unfair. (Victoria, School #2, Grade 4, 5/17/11)

Victoria's quote emphasizes that teachers are engaging in instructional practices that have more to do with teaching reading skills to enhance test performance rather than teaching students literacy for life. When the test is used inappropriately, instruction that is not valuable to truly develop students' literacy ensues. Not only is the test not designed for emergent bilingual students, but also the use of the exam results is also inappropriate for these students.

Teachers have four basic critiques of the ELA assessment data with respect to emergent bilingual students:

- 1) test construction confounds language modalities and performance indicators do not match what they report to assess;
- 2) the ways in which data is reported to teachers is virtually meaningless because the data reflects a range of "tangled" skills;
- 3) some emergent bilingual students lack the proficiency levels in English and testing skills in order to be properly assessed in English reading with this assessment;
- 4) test data leads to test-prep which teachers believe does not contribute to students' reading skills.

Their critiques mirror the criticisms of this type of assessment data offered in the literature review (Butler & Stevens, 2001; Abedi et al. 2006; Abedi, 2008). Also, as Popham contends, the

use of summative data to shape instructional practices is an inappropriate one (2008). For these reasons, the utility of summative data to contribute to teacher knowledge and to drive instructional practices suited to emergent bilingual student learning is extremely limited.

While these arguments are sound, they do not change the course of how this data is expected to be used by teachers by school administrators and ultimately policy-makers. This mismatch between the appropriateness of the test for emergent bilingual students and the expectation that these results will indeed flow into instructional practices leads to a conflicted relationship between teacher knowledge and the ELA assessment of emergent bilinguals. This awkward relationship between ELA results and teacher knowledge will be described in the following section.

### **Operating Instructions: Knowledge and the ELA Exam**

Although conceding that they learned from the ELA, teachers did not provide specifics about *what* they learned from the results of this assessment. Data analysis revealed that teachers felt that they should be learning or were expected to learn from these results. One teacher voiced this very clearly when he assured me, “We are definitely learning from the test, it’s not that we are not learning” (Sameer, School #4, Grade 4, 12/1/10). Yet, this teacher did not describe the knowledge he derived from the assessment. In general, teachers interviewed expressed that they felt they *should* learn from assessment results, but in actuality what they reported they did with these results indicated the opposite. An analysis of responses regarding the types of knowledge that teachers gleaned from assessment, points to the finding that teachers did not gain knowledge from the analysis of summative data of emergent bilinguals.

This section continues to make the link between teachers' critiques of the ELA and their mistrust of this data which in turn leads them to disregard of data. I also provide evidence that much of what teachers do with data comes from ritualized assessment practices. As teachers "go through the motions" of examining assessment data, instead of learning about their students, teachers *learn* to operate within a ritualized set of assessment practices that are sponsored by school administration. These ritualized assessment practices are a form of operating instructions that teachers are required to do when analyzing assessment data. It is important to note that schools sponsor the specifics of ritualized assessment practices are framed by test-based accountability measures. *Test-based accountability* narrowly defines success in terms of *school test performance*, and thereby also narrows the focus of school leaders to ensure that teachers are engaged in data-driven instruction, regardless of its appropriateness (Menken, 2008; Linn, 2003; Elmore, 2005). Ritualized assessment practices give the illusion that teachers are deeply engaged in analysis of summative assessment results — when in fact, what they are doing is using the test results to identify patterns of student performance and sort students into instructional groups in ways that may not be helpful. Lastly, I also provide evidence that rather than the summative assessment data providing teachers with knowledge of students, it is their existing knowledge of students, acquired through classroom interaction that is used to make sense of data from summative assessments, rather than the other way around.

As described in the previous section, teachers harbored deep concerns about the ability of this assessment to capture the reading abilities of emergent bilinguals. Since teachers do not believe that the test data provides insight into emergent bilingual students' reading abilities, they do not actively use this data to learn about their students. As the quote below illustrates, teachers

remarked that the ELA results for emergent bilinguals do not give a clear picture of student reading abilities and therefore, teachers do not refer to them to learn about their students. In the following quote, Kai expresses not consulting the ELA data at all. In the excerpts from interviews, emergent bilingual students are referred to as “ELLs.”

Laura: What about large-scale data? Do you use the ELA to assess your ELL readers?

Kai: Not so much. Not really. Honestly I don't [consult this data]. The scores sometimes are provided to me in a packet and I put them in the assessment binder. (School #2, Grade 4, 11/14/10)

This fourth grade teacher candidly describes what most teachers do with assessment data derived from the ELA. They do not consult this data to learn about their emergent bilinguals' reading development. Instead, this data is filed away in assessment binders and not actively referred to by teachers to learn about their students. Teachers cannot learn from data they do not actively consult. Her quote also demonstrates one step in the ritualized assessment practices supported by schools. Teachers are all provided with test results and an assessment binder, a repository for class assessments. This documentation, while usually not referenced frequently by teachers, provides evidence and the illusion that assessment data is organized and maintained by teachers in the case that school administration or municipal officials inquire about data use. Kai's quote demonstrates that teachers *learn* to maintain records of supposed data use; that is, what they do learn well is to follow ritualized assessment practices required by the school culture.

Teachers have been made to believe that they should examine the data for trends that would reveal valuable information about student performance. In the following quote, Beatriz demonstrates the conflict about using this data. Although she says that the process is helpful, in

fact, what she is really saying is that she learns nothing new about her students from identifying patterns from performance levels of her students. She expresses it this way:

I think it's interesting to look at the numeric score and to see who are the students who scored below a certain level and to see whether or not, they are ELLs. Unfortunately most of them are and most of them are students with IEPs. (Beatriz, School #1, Grade 4, 11/26/10).

Beatriz' internal conflict is evident — although she states that the data is “interesting,” it reveals nothing new, instead it just confirms that the lowest performers are the ones expected, those with disabilities. This internal conflict in actuality stems from the external expectation that the summative data will yield valuable information about students when in fact, it does not. Beatriz' conflicted response stems from the prioritization and expectation of using summative data to impact instruction which ultimately is the consequence of a test-based accountability system. Beatriz has learned to think that examining the ELA for trends will be valuable and thought-provoking, while her reality indicates the opposite.

While some teachers are aware that the data, for reasons enumerated previously, does not provide an opportunity for teachers to acquire knowledge, some suggested how the data could be more helpful to them and hence be a source of learning. Suzie, a fourth grade teacher, advocates that teachers could acquire knowledge about emergent bilinguals' reading development from an analysis of individual items on the test. Alongside her critique of the reporting of performance scores, she suggests:

I think that having just the score, 1, 2, 3 or 4, doesn't say very much. But I do think that the type of questions, if I was able to have a test in hand and look at the way that my students answered, that would give me useful information. (Suzie, School #2, Grade 4, 11/12/10)

Suzie suggests a process of examining student responses alongside test questions. This process would be an opportunity for teachers to learn about their students and the exam, thereby building both student knowledge and assessment knowledge. Although many teachers expressed their dissatisfaction with the usefulness of how the scores were reported, they did not describe their engagement in this kind of “tandem analysis” in which analysis of test items wording is done alongside the skills students are expected to have to answer the question. Instead, teachers repeatedly expressed that the ELA assessment results were used both by school administration and teachers to identify trends within their class across their students, without taking into account the particularities of the student population. This analysis was comprised of the identification of patterns by sorting students according to performance levels. For example, all students who obtained “2s” would be grouped in one enrichment group, all the students who received “3s” in another and so on. The point of Suzie’s quote is that although teachers are aware of what would help them to learn about their emergent bilingual students and the assessment process, these processes are not included in ritualized assessment practices in which they are engaged as a result of the test accountability culture.

In the following quote, Beth, a third grade teacher, speaks about how knowledge about a student’s total learning profile, which includes their reading abilities, prior knowledge of test-taking, and learning style, may impact performance on the test. Beth’s narrative is an example of how teachers are uniquely aware of students as a total entity — students have unique needs depending on their personality, history, and language development, among other factors. Like many teachers, she claims that the test is not an accurate measurement of reading abilities for many students because of how the test is constructed. She also voiced serious concerns about

whether the ELA results had any merit for emergent bilingual students. However, most importantly, she demonstrates how she employs *pre-existing* student knowledge acquired through classroom interaction in order to identify the factors that affect her emergent bilingual students' test performance. According to Beth:

Kids, who are strong readers, whether they are first or second language learners, do better than poor readers do on the test. I'm not sure it gives us a whole lot of more information than that. When you have 5 different reading selections, how much is the score based on what they are able to do, their languages, what their mood is that day or if they are able to sit still that long? I think particularly when the teacher knows that the student had challenges for whatever reason, and it's [the test] at a higher level than what the kid is able to do, then it's an absolutely useless assessment. (Beth, School #1, Grade 3, 11/4/10)

As Beth names the range of factors — reading selections, language proficiency, socio-emotional state, and reading level — she demonstrates the use of knowledge acquired from other realms of her teaching practice in order to interpret this data, rather than constructing new knowledge from the process of analyzing the ELA.

What kinds of knowledge do teachers acquire from the analysis of the ELA results of emergent bilinguals? Throughout the 20 interviews conducted, teachers did not report acquiring any of the four types of teacher knowledge (pedagogical content, student, assessment and bilingual) from examining the results of the ELA. Rather teachers are learning to use operating instructions or *ritualized assessment practices* which allow them to go through the motions of analyzing summative data, when in reality this data is inappropriate for use with emergent bilinguals and yields limited information about their reading development. Furthermore, in interpreting assessment data teachers draw from knowledge they already possess as a result of their teaching practices and classroom interactions with students. The knowledge they

demonstrate about either students or the assessment is not acquired through analysis of ELA data. Teachers, like Beth, use the knowledge they hold about emergent bilingual students to bolster their reasons why the ELA does not accurately capture the reading development of these students.

This summative assessment provides little new information about emergent bilinguals. Teachers are not learning about their students from it. Yet, they are learning to work within a ritualized set of assessment practices which give the illusion of deep consideration of this type of data. And yet, the test is often consulted and analyzed for trends by teachers, as mandated by school administration and framed by test-based accountability measures. How and what teachers learn about how to use this assessment to shape their teaching of reading is influenced much more by ritualized assessment practices, rather than by an analysis of the assessment data itself. The next section looks at how the ELA impacts teacher practices and consequently instruction for emergent bilinguals.

### **Mining the ELA for Instruction: “We are Using What You Can Use”**

Although teachers do not gain new knowledge about their emergent bilingual students from the ELA, they convey that the ELA does impact their instruction. Teachers learn how to operate within ritualized assessment practices which provide them with steps about how to “analyze” ELA assessment data, irrespective of the student population, and following a test-accountability culture. These ritualized assessment practices take the place of teacher knowledge. Ritualized assessment practices consist of using students’ performance levels to identify patterns and group students regardless of whether their scores relate to language proficiency or reading difficulties or both. As such, these ritualized assessment practices by

teachers do not reflect adaptive practices that take into consideration the instructional needs of emergent bilingual students. These ritualized practices of teachers result in two teacher actions – 1) the identification of patterns of performance across students, and 2) the grouping of students by performance levels.

Zoila, a third grade teacher, describes the assessment performance at her school. In the following quote, she describes that she, along with other teachers, examines trends by identifying patterns of students who perform at a specific performance level. The motivation behind this type of analysis is that it will drive instruction. Zoila says:

What we do as a team is we look at the patterns from the previous year, for other third graders and yes, we look at trends and the things that they didn't get, so we can perhaps modify our curriculum a little bit. (Zoila, School #2, Grade 3, 11/12/10)

Another teacher at School 2, Aura, details how ritualized assessment practices occur at all levels – within and across grades – and how school resources support it. Her quote also describes how this analysis of looking at trends directly informs what is taught in her classroom. Aura says:

We look at the ELA – school-wide and grade-wide – to see where our strengths are and where we need to improve. There is a direct connection between the ELA and what we do with our reading specialist. We meet with her once a week and then that trickles down to our planning because we plan based on the discussions what we have and the information that we get from her. We talk about the fact that we need to concentrate more on this particular section or this particular area. Then we interject that into our planning. (Aura, School #2, Grade 3, 11/12/10)

Aura's quote highlights the role of school personnel in supporting the ritualized assessment practices, thus underscoring how seriously the narrative that ELA data can be used formatively is taken at most schools. Furthermore, as described by teachers, often this is the only type of help

that schools provide in analyzing any type of data, which is indicative of how priorities set by test-based accountability measures misdirects school resources.

Sameer, a fourth grade teacher at another school, explains how this process is translated directly into classroom practices. He describes how performance indicators for all students, regardless if they are emergent bilingual students or not, informs instructional priorities. Sameer states:

We looked at this data as a grade: which questions did they score the lowest from third grade? And we were trying to see if this question has to do with referring to a chart or summarizing a paragraph. Based on that analysis, we said, ‘let’s spend this period focused on summarizing or retelling because we need it.’ (Sameer, School #4, Grade 4, 12/1/10)

These quotes together demonstrate how teachers identify patterns based on performance indicators and how these patterns impact what teachers teach in the classroom. Because these behaviors do not usually lead to blatant test-prep sessions (where students are working on practice tests), many teachers do not believe that they are teaching to the test. However, because the accuracy of these results for emergent bilingual students are in question, this type of instruction can be considered test-prep because it is focused on teaching students skills to do well on the test, rather than being a true measure of what emergent bilingual students need to do to be successful readers.

Another product of teachers’ ritualized assessment practices is grouping. After identifying trends in order to shape instruction, teachers often meet with instructional support staff — professional developers, coaches and other grade-level teachers — to use the ELA scores to sort their classes into small groups to receive targeted instruction. Suzie, a fourth grade

teacher, reflects on this process that many teachers described in analyzing student data. She explains how the analysis of performance levels is used to construct groups:

So going into a new year, one of the first things we do we'll look at their ELA and their math scores and kind of like see the breakdown of their scores and see what areas they need help in. What's the next level? If they were a high two or a low two, and also based on the ELA analysis from the previous year, the reading specialist makes reading intervention groups where the students are pulled out for reading intervention usually twice a week for each group and I believe that's done all for English. (Suzie, School #2, Grade 4, 11/12/11)

This quote is representative of how teachers work the data from the ELA and possibly other large-scale test data. Teachers are essentially supported to use proficiency levels to track students into instructional groups. Additionally, the instruction that occurs for these groups is prescribed by their performance level – a rather imprecise measure – on this large-scale literacy exam. Furthermore, the above quote also suggests that there is a bias towards the analysis of students' literacy abilities in English (and not in languages other than English) and in the allocation of enrichment services that are available to students.<sup>25</sup> It is important to be reminded that all teachers interviewed teach in Spanish-English, dual language bilingual programs. Therefore, enrichment services should be offered in both languages and not solely in English in order to impact the reading development of emergent bilingual students.

The effect that these ritualized assessment practices have on emergent bilinguals is especially troubling. Those emergent bilingual students whose English proficiency levels are low are often grouped with students who perform poorly on large-scale standardized tests for

---

<sup>25</sup> In New York City, enrichment services, usually provided to students during extended day (37.5 minutes for small group work as stipulated by the teacher contract), consist of small group instruction. Students are grouped by ability homogeneously and provided with support based on their performance level from the ELA.

reasons other than English proficiency. This composite group of struggling students is instructed in English only, despite being in dual language bilingual schools. Lastly, the instruction that takes place during this instruction is based on what students are perceived to need in order to raise achievement on the *test*, rather than reflecting a true evaluation of their reading needs. This kind of instruction therefore may be unsuccessful at truly raising the reading abilities of emergent bilingual students. Therefore, the types of instructional practices that are implemented as a result of the ELA analysis are prescribed by the testing demands of the test.

These ritualized assessment practices described by teachers at all schools were remarkably similar to each other. Teachers identify patterns of performance and sorted students into groups based on performance levels. This similarity was characterized by the process' unidirectionality. The assessment results trickled down to students, as teachers identified patterns within and across grades ultimately grouping students by performance level without regard to student-level factors, which may better inform this process. This is a limitation of data-driven reform, in which students are merely placed within the "flow" of assessment data and their unique needs are not considered. The widespread use by teachers of similar ritualized assessment practices at all four focal schools points to the reality that these school-sponsored assessment practices stem from the test-based accountability system that has been set in place under NCLB. As described in the literature review, test-based accountability systems have been critiqued for the narrowness in vision and their potential to lead to practices which instead of improving educational outcomes lead teachers to "slim-down" and focus their curriculum on what is tested.

Most importantly, the teachers' ritualized assessment practices do not take into account emergent bilingual students, although all teachers interviewed taught at dual language bilingual schools. Although dual language bilingual schools are considered a strong form of bilingual education with goals of bilingualism and biliteracy, in the case of assessment, teachers continue to use this data in ways they know are inappropriate. In fact, although teachers spoke about the use of ELA assessment data to identify patterns of performance and sort students into groups, few of the teachers mentioned consideration of emergent bilingual students. This lack of regard for the assessment needs of emergent bilinguals lends evidence to the argument that these practices form part of an overriding ritualized assessment practice that does not result in the development of knowledge that is of any consequence for emergent bilingual student learning.

The identification of patterns of performance and grouping of students according to performance levels from the ELA lead to different instructional practices. And yet, the types of instructional practices that are implemented as a result of the ELA analysis have more to do with fixed practices at the school rather than an understanding of emergent bilingual students; they are *ritualized assessment practices* or static methods of working with data that do not respond to the unique characteristics of students, especially emergent bilingual students.

Sameer, a grade four teacher, shows how the identification of patterns of student performance directly connects to how students can achieve higher scores on the test. Sameer describes:

The scores are a big deal. We try to see what they did to get that score and using that to see how we can get them to the next score. We focus on the behaviors and strategies and skills to get that 3 or if you have a 3 to get them to that 4. So that's the main thing that I try to focus on: what are they missing to get to get to the next step. (Sameer, School #4, Grade 4, 12/1/10)

As a result, the instructional practices that are implemented are not attuned to the developing reading needs of emergent bilingual students. Again, emergent bilinguals are not explicitly mentioned by teachers in discussions about how performance levels are used to shape instructional priorities. This lack of regard for emergent bilinguals students in shaping instructional practices as a result of assessment data, reflects the negligible impact and influence that teachers' store of bilingual knowledge has on these assessment practices.

In fact, the most frequent instructional practices implemented as a result of consulting ELA data do not reflect complex ideas of either reading or bilingualism. The instructional practices most frequently implemented were the teaching of the main idea, vocabulary, and context clues. These three skills, although helpful for any reader, are just the tip of the iceberg in the formation of a competent reader. Furthermore, these reading skills are often taught with the purpose of test-taking rather than becoming a competent reader. One third grade teacher states;

We talk about what's a main idea and how we can find that. When they [in the test] say main idea, what does that mean, that's what we are looking for [in our students]? We don't necessarily talk about those things in reading. (Pamela, School#1, Grade 3, 11/4/11)

Pamela makes the distinction between teaching a skill in order to do well on the test and teaching students skills and strategies to be successful, life-long reader. This is yet another example in which the data is used for test-prep although disguised as data-driven instruction.

Furthermore, teachers analyze the ELA not only for the purpose of isolating which reading skills students have yet to acquire, but also to identify testing strategies that have little to do with developing into a successful reader. Judy, a veteran fifth grade teacher concedes:

What we are noticing is that the answers too many of the questions are hidden within the text in weird places. So they [students] sort of have to go back and reread the whole thing to find one answer, which is becoming harder for the kids. We took the test apart and we saw that some answers were sort in tricky places. So we have to teach them new strategies – new places where you can find the answers. It's like a lot of test-taking instead of becoming a reader. (Judy, School #4, Grade 5, 12/1/10)

Judy's quote demonstrates that instead of building pedagogical content knowledge or an understanding of how to teach reading more effectively to emergent bilingual students, teachers are studying the test structure and its passages in order to understand how to teach students to answer the test questions. In other words, the test is being "hacked" so that teachers can ensure that instruction of reading skills matches testing demands. Therefore, although instructional practices do emerge from examining the assessment data, it is questionable if this instruction holds potential for the teaching of reading for emergent bilingual students.

Instead of acquiring assessment knowledge, teachers have learned how to operate within fixed and ritualized assessment practices. These school-sponsored behaviors function as operating instructions. However, instead of serving as a protocol to aid teachers in developing knowledge about their emergent bilingual students or even about the assessment itself, teachers are going through the motions and using the data to group and sort students according using performance levels. This process is particularly imprudent for emergent bilingual students because their score on this exam does not distinguish between their language proficiency and their reading ability. These practices become even more worrisome given that teachers are aware that the results of this test are not valid for emergent bilinguals. It reveals that teachers are caught within a process that does not make sense to them.

The next section highlights a promising practice of using the ELA with students that two of the teachers interviewed discussed. It is featured here in order to demonstrate how the assessment performance can be altered in order to yield teacher knowledge or instruction that is meaningful or beneficial to emergent bilingual students.

### **Involving Students: An Opportunity to Rescue ELA Data**

A few teachers laid out some promising proposals that could inject the process of analyzing ELA data with more purpose for emergent bilingual students. These adaptations present promising opportunities to alter these fixed practices which do not account for the assessment needs of emergent bilinguals and create spaces for teachers to apply their knowledge of bilingualism to the analysis of test results. One promising practice reported by one teacher at School #2, was referred to as “Data Day.” During “Data Day” teachers work individually with students to interpret test results and help students set goals. This practice may have the potential to alter the process of analyzing summative data for teachers because it, hopefully, interrupts the unidirectional flow between the ELA data and students. Students can explain and describe their experiences in taking the exam, and in this way illuminate for teachers the undeniable interplay between the test and student abilities. Aura gives the following description of Data Day:

With third grade on, which are the testing grades, we do something called ‘Data Day.’ We get the results back and we take a day to give each student their results. So we have them do a project that will keep them occupied while we talk to each student one by one and show them the results on the test, talk about what it means and help them to create goals for themselves. It’s great! (Aura, School #2, Grade 3, 11/12/10)

A similar process that involves students (that is not currently implemented) was described by Susanna. This practice, similar to the one stated above, would also involve students in interpreting their test data to set goals. Susanna describes this process:

Sometimes we find our kids getting really caught up in the number, I'm a 3. You are not a number. Describe your strengths and weaknesses, so that we can move you forward, is what I'd like to really move toward. It would be really great to get into that with students – what particular aspects of the test they did well on and what particular aspects of they need to work on. So that they stop seeing things as just numbers and start seeing things specifically. These are goals. These are not things we do. (Susanna, School #3, Coach, 11/26/10)

This call by teachers to involve students in reflecting on their performance on the ELA can also be considered an appeal to teachers to examine the data alongside the emergent bilingual students, including their bilingual knowledge. This process of inserting students into the conversation about assessment data may allow for teachers to make the process of consulting ELA valuable by inserting a formative process into it, rather than remaining locked within generic assessment practices. The danger of these practices would be to just inform students of their performance levels, rather than involving them in revealing to teachers their experience of the test. In other words, in order for this practice to help teachers gain knowledge of their emergent bilinguals, a formative process must be inserted into this summative assessment.

### **Learning and Not Learning from the ELA: A Summary**

In sum, although large-scale standardized assessments are used to describe trends within the student population, the trends identified often track and misidentify the instructional needs of emergent bilingual students. Data from the interviews indicate that teachers are aware that this assessment provides very little information about emergent bilingual students and how to instruct them in literacy. Furthermore, when the results of large-scale standardized exams do in fact shape instructional priorities, the instruction that emerges from the analysis may be misguided for these students.

Therefore, summative assessments do not provide a solid platform for teacher learning about emergent bilingual students' reading development. Rather, their use is characterized by school-sponsored ritualized assessment practices which facilitate the implementation of instructional practices with questionable merit for emergent bilinguals. One of the principal reasons for this lack of fit between the ELA and both teacher knowledge and instructional practices is that this summative assessment is being used formatively, although it was constructed as a summative assessment. That is, through ritualized assessment practices, a test which is designed to provide benchmark results is instead being analyzed through the identification of patterns of performance and grouping, and hence used to shape instruction, when this use is both inaccurate and potentially harmful to emergent bilingual students. The underlying reason that ritualized assessment practices are implemented across these dual language bilingual schools is because test-based accountability measures place enormous pressures on all school leaders to demonstrate that they are using the results of these tests to shape instruction.

The following are findings regarding teachers' work with the ELA data of emergent bilinguals:

1. After consulting ELA data, teachers did not report any gains in knowledge of their emergent bilingual students. They voiced a number of concerns about the validity of exam results for these students. These concerns matched those found in the literature.
2. Teachers' use of assessment data was characterized by ritualized assessment practices; that is, fixed behaviors sponsored by school administrators that did not take into account student characteristics or teachers' knowledge of students derived from their classroom interaction.

3. The main instructional practice that was implemented as a result of ELA assessment practices was grouping. This is a potentially troublesome outcome for emergent bilingual students whose low English proficiency may lead them to be placed in groups that do not reflect their reading ability, and thus not match their instructional needs.
4. Data from the ELA was used formatively. Since the ELA is a summative assessment, it is an inadequate use of this type of data.

While teachers spoke about the constraints of summative data, often they mentioned the potential of formative data. I now turn to the analysis of formative assessment in order to examine if the use of this type of data matches teachers' expectation of it for informing their understandings of emergent bilingual students.

## **Part 2: Teachers Speak about Running Records, a Formative Assessment**

Given that large-scaled standardized exams are both limited in their ability to give teachers information about their emergent bilingual students and to meaningfully aid in shaping instruction, I now turn to what teachers of emergent bilinguals learn from the process of administering and analyzing data from running records and how this learning may be translated into instructional practices.

Running records differ considerably from large-scale standardized exams. First and foremost they are intended to be formative assessments. As such, they are the first step of the formative process of administering, analyzing and tailoring instruction for students. They are administered more frequently than large-scale standardized tests. In schools where the teachers who were interviewed for this study worked, they were administered from three to five times a year.

In comparison to large-scale standardized exams, running records are focused more fully on one specific aspect of literacy – reading. Running records are conducted when a teacher documents and analyzes a student’s reading for the purpose of identifying the skills, strategies, and level that a student employs and reads at. Running records can be formal or informal. This dissertation focuses on formal running records. Running record kits contain a wide range of leveled texts that are manufactured by a variety of publishers. Schools elect to purchase a particular kit depending on their preferences for a given product. New York City schools, as previously stated, are required to demonstrate a collection of on-going data about their students. As such, all schools in New York City employ some form of running record assessment. In administering formal running records teachers use level texts along with a documentation sheet (see Appendix 5). Before reading the text, teachers usually gather information about students’ prior knowledge about the topic of the text. Afterwards, while students read, teachers take detailed notes about students’ errors, attempts at self-correction, and fluency and phrasing. Following the reading, teachers ask students comprehension questions. All this data is used primarily for two purposes: 1) to identify a student reading level, 2) to analyze student miscues.

There are many leveling systems that are used to describe students’ ability to read at a given difficulty level. The Fountas & Pinnell leveling system is the one used most frequently in New York City. Leveling systems provide a description of text difficulty. Furthermore, most leveling systems are pegged to grade level standards. The Fountas & Pinnell leveling system spans from A to Z. Students who read at level A are emergent readers who are beginning to connect meaning to print. Students at level Z are fluent readers. A detailed description of Fountas & Pinnell levels and grade level correlations are found in Appendices 8 & 9.

Running record assessments assess the essential components of student reading such as decoding, fluency, phrasing, and comprehension on an individual basis, although students must demonstrate their knowledge and skills orally. This exclusive focus on reading behaviors and skills is in contrast to the large-scale standardized assessments which assess reading through listening and writing. At the heart of running records is teacher documentation of student reading fluency, errors, and comprehension. This documentation serves as data for miscue analysis, a fundamental aspect of the way in which running records can yield information about students. In miscue analysis, students' errors are analyzed in order to understand how students' think and employ strategies to read. Miscue analysis is the core of this formative assessment as it requires teacher interpretation and provides teachers with a detailed, qualitative understanding of student reading. For example, if a student reads "My mother *makes* cookies" when the text reads, "My mother *bakes* cookies," it would provide the teacher with specific information about this student's reading skills. First, the student substituted "make" for "bake," therefore preserving meaning of the sentence. The student also possibly looked at the word ending, but did not consider the initial sound of the word. This may clue the teacher into the information that the student needs to pay greater attention to word beginnings. Finally, the administration of formal running records provides teachers with a student reading level and indicators of their comprehension levels. Given all of these reasons, running records are often considered by teachers to provide a formative assessment opportunity that is better able to capture the reading abilities of emergent bilingual students.

However, all teachers may not conduct both parts of this analysis – miscue analysis and identification of the student reading level. Some teachers may do a miscue analysis solely

because they know a particular student is not reading at a higher level. Other teachers may only identify a reading level. In this case, not conducting the miscue analysis is a missed occasion to know more specific information about students' reading.

Teachers spoke at length about how they learn and use the data from running records for emergent bilinguals. All teachers who were interviewed worked in dual language bilingual programs and therefore administered and consulted running records in both English and Spanish. This in itself is different than with the ELA discussed above which is given only in English. The first section in this part of the chapter will outline findings about how teachers acquire knowledge of emergent bilingual students as readers from running records, as well as the contextual factors that influence it. In doing so, assessment knowledge will be explored along with the types of adaptations that teachers make of running records based on their familiarity with this assessment. Bilingual knowledge will also be explored in depth paying particular attention to how teachers manage data from both English and Spanish running records. The following section will discuss the instructional practices that teachers report implementing after consulting reading record data of their emergent bilingual students. The section will conclude by laying out teachers' proposals to make the use of running record data more effective for emergent bilinguals. I turn now to examining teachers' responses about student knowledge acquired via running records.

### **Seeing Students as Readers: Student Knowledge Acquisition from Running Records**

Teachers overwhelmingly stated that they learned through running records. Unlike the dissatisfaction surrounding the ELA, there was little criticism of running records. Above all, teachers noted gaining student knowledge through these formative assessments. Student

knowledge or the understanding of the skills and knowledge possessed by students was described in many ways by teachers. Time and time again, teachers identified what they learned about their students' reading abilities *through* administering the assessment. This is an important difference with the ELA discussed previously. One teacher talked about how the data from running records afforded her with the ability to focus on the habits, skills, and thinking processes in reading that her students possess. Judy, a fifth grade teacher at School #4, remarks that through running records she is able to learn about her students' reading in context. This quote is representative of how a majority of teachers spoke about what they learned about their students from this type of assessment. She states:

They [running records] sort of give us a very good indication of what type of reader the child is. You can tell their habits as a reader. You can tell their fluency, their vocabulary, their comprehension. You also get a picture of how they process the information. (Judy, School #4, Grade 5, 12/1/10)

The types of things that this teacher describes learning about her students through running records match the intended purpose of this assessment in general. She, like other teachers, learns about the strengths and weaknesses of her students' reading abilities through considering the skill level of each student along these essential components of reading.

Another teacher, Kai, describes the types of questions she asks about her students' reading as she administers the assessment. Her response reveals that the type of student knowledge that teachers acquire is both about students' skill level as well as their reading behaviors. The following quote is also illustrative of the majority of teachers' responses about what they learn from this type of assessment:

The purpose of the running records is to assess fluency and reading strategies. So are students using syllabification to figure out a word? Are they breaking the word down and

sounding it out phonetically? Are students using the picture in order to help them decode a word? So we code all of their errors and the coding of those errors helps me know whether or not they are actually using those strategies. (Kai, School #2, Grade 4, 11/14/10)

The above quote describes the types of behaviors that teachers are looking for when they assess their students. It also demonstrates how the *process of administering a running record* is equally important to obtaining student reading level, often considered the final step in completing this assessment. It is these observations of student reading that provide teachers with a full picture of students as readers.

Although teachers described learning about student reading processes during the administration of the running record, few teachers talked about using miscue analysis as a way to examine the data afterwards – although it is key in analyzing running records. The following teacher’s description of how he makes sense of running record data may provide a glimpse into why many teachers do not engage in miscue analysis of their running record data:

Well the running records, after they are done, there are certain components that the child has to master in a way. So they have to have 96% and above of miscues [total percentage of word read correctly], they have to read 100-150 words per minute. So there are criteria that a student has to go through in a passage regardless if a child passes that criteria or not, you see this child is doing great on fluency but he’s lacking in inferring. So the next time you do an individual conference you are not going to tackle fluency, you are going to tackle inferring. So you get to see individually what each child is lacking regardless if they pass the level or not. (Sameer, School #4, Grade 4, 12/1/10)

Sameer’s narrative suggests that teachers may hold misconceptions of miscue analysis. They may view miscue analysis solely as the percentage of errors that students commit rather than as an inquiry into understanding the type and quality of these errors as described previously.

Only one teacher in the sample spoke at length about how she uses miscue analysis in order to learn about how students' reading and language skills develop in tandem. This singular quote is presented to demonstrate the potential that this type of analysis can have on understanding students' bilingual reading development. Aura describes:

It [miscue analysis] can tell me how they grow over time and what kind of job I'm doing teaching them the strategies based on the analysis of errors. If the errors start out as visual errors, which what we expect it to be if they are learning the language and learning the sounds, then visual errors are probably what we would see most. If they are going away from visual errors into other types of errors, then we know that their language is growing and that their reading ability is growing as well. (Aura, School #2, Grade 3, 11/12/10)

Aura's quote demonstrates the integrated knowledge of reading and language development that can emerge from an analysis of the students' miscues. Although it can be argued that she already possessed the assessment knowledge of how to conduct miscue analysis with an eye on language development, her perspective is still important. It demonstrates how putting both assessment and bilingual knowledge to use can yield valuable information about students' simultaneous reading and language development. Lastly this quote focuses on the interconnectedness between different types of knowledge. If teachers can apply assessment knowledge and bilingual knowledge while they are evaluating students, they will glean more accurate knowledge about their emergent bilingual students' reading development.

Although these findings demonstrate that teachers are learning important information about their students' growth in reading, they also show that teachers administer and analyze running records generically; that is, without attention to whether students are English monolinguals or emergent bilinguals. As a general process, running records do not provide

specific information about the development of emergent bilingual students' reading. The ways that dual language bilingual teachers in this sample currently use running record assessments with emergent bilingual students are the same they use with monolingual readers. Teachers are attuned to student reading skills and their complementary reading behaviors, but they do not analyze them to specifically understand how their emergent bilingual students' reading and language development are intertwined.

Analyzing running record data in English and in Spanish to develop an integrated understanding of students' bilingual reading development is a complex endeavor. And yet, the reasons why teachers do not use this assessment to yield student knowledge that is specific to bilingualism has much to do with the culture of "ritualized assessment practices" described in the findings about the ELA, and will be discussed in the next section.

### **"I do what I am Told to Do:" Going through the Motions in Administering Running Records**

An initial question that these findings pose is: Why does the acquisition of student knowledge stop short of providing knowledge of students that is specific to their bilingualism? Part of the answer may lie in teachers' understanding of the purpose of running records, coupled with contextual factors that frame the use of running records in schools. Teachers often gave a standard answer when asked the question, What is the purpose of running records? The response more times than not included the explicit mention of identifying students' reading levels regardless of the language the running record was taken in. As previously described, the student reading level is solely one piece of data collected from the running record.

Constanza's response to this question is emblematic of how nearly all of the dual language bilingual teachers viewed running records. When asked about the purpose of running records, she replies:

The purpose of the running record is to see – a. what level they need to be at and b. to see what books they need to read at an independent level and to see if they are the stage where they need to be at this particular month or the following month and progressing to meet standards by June. (Constanza, School #3, Grade 5, 11/26/10)

This quote demonstrates teachers' unbalanced focus on reading levels as compared to the other data collected from the running record such as students' miscues. One reason teachers concentrate on this aspect of the running record over others is that these levels must be reported to school administration, a consequence of the test-based accountability climate. Thus, running records take on the character of a high-stakes assessment and are often constrained by the same "ritualized assessment practices" that are part of the school culture. The joint pressures of school mandates and the time required to assess all students in two languages creates for many teachers a situation in which teachers are completing the assessments in order to meet the assessment requirements set by the school, rather than to gain insight into emergent bilingual students' reading abilities. The impact of these factors on the administration and analysis of this assessment cannot be underestimated.

School principals require teachers to report students' reading levels in both languages several times a year. Therefore, teachers are motivated, in part, to conduct these assessments because of a school-level mandate, rather than from a personal desire to gain a greater understanding of student reading. Jaime very candidly expresses why she conducts running records:

Pretty much honestly, I do what I am told to do. I learned about running records when I was in college over 10 to 15 years ago and it's what I've been using since the beginning and what I feel works for me. They've become more and more informal because I think I can gather more information without having a script in front of me and that's because I'm experienced. (Jaime, School #1, Grade 3, 11/26/10)

Additionally, administering running records and analyzing the results requires time. This is time that is not officially allocated to teachers by schools. Teachers generally administer running records during students' independent reading period, which is an instructional period. The analysis is done during teachers' preparation period or after school. This "time crunch" to assess and analyze the assessment data was felt and voiced by a large portion of the teachers interviewed. The following fifth grade teacher, Amelia, describes the time demands that running records present, in particular the analysis. In her specific case, because she works with a co-teacher, they are able to discuss assessment results informally. She asserts:

It's not like we take all of our assessments and we say now we need to figure this out. We sort of informally say this is what happened, the assessments take a long time. So it could be a few weeks and then she might be doing assessments. So sometimes we are both assessing and we talk about it, we get together either at lunch or during our prep and we talk about what is going on and we sort of talk about the concerns. We see, oh wow, what is going on with this kid and we talk about this kid is really good and he moved up a lot. There are certain kids that stick out and we sort of come back and forth and talk about it. But it's not like we have a whole block of time to do this, we just don't. (Amelia, School #5, Grade 5, 12/1/10)

In fact, most teachers did not even go this far as to discuss the progress of their students as evidenced by the running records. For instance, Kai states that "a great deal of the stress I have is I'm going to have that piece of paper [with assessment results]; whether I look at it, whether I analyze it, is another story" (School #2, Grade 4, 11/14/10). This pressure to "hand in" reading levels from running records to school administration may also shape teachers'

perceptions of the purpose of this assessment, as well as the extent to which they analyze the information collected from running records. Furthermore, the time pressure becomes compounded for dual language bilingual teachers who must administer and analyze running record data for each student in two languages.

Teachers stated that running record assessment cuts down on the time they have to teach. This creates an antagonistic relationship between assessment and instruction for teachers, exactly the opposite of what formative assessments intend. Constanza concedes:

It takes a long time to gather, a really long time to read through it thoroughly and understand what it really means, what are next steps are – how do we group? So we are kind of overwhelmed by all the data and we are like, the deadline and we got to do this and we got to do that and report cards and by the time we gather all the data and understand it, it's already time to assess again. There is more testing than teaching. It takes so much away from your teaching and we are battling that a lot. (Constanza, School #2, Grade 5, 11/26/10)

This viewpoint that running record assessments are eating away at instructional time was common. In fact, many teachers expressed resentment that these assessments “interrupted” instruction. Only after assessments were completed could teaching resume.

In addition to the impression that the running records temporarily take away from instructional time, the pressures of reporting requirements established by schools may steer teachers away from the formative character of this assessment. Due to the aforementioned pressures, many teachers focus solely on finishing the assessments and reporting the results. The results are then filed away. In fact, teachers feel the prescribed “assessment window” does not coincide on a timeline that matches students’ needs. In the following quote Marisol, a fourth

grade teacher, speaks about how the required assessment period may not match the students' progress. Marisol argues:

Timing is an issue, especially in dual language bilingual classes. You know we could argue for a little bit of extra time. Sometimes you know that the student is not ready for the next level, but we have to do it, so it's a waste of time. (School #4, Grade 4, 12/1/10)

The quote above highlights that the rush to complete assessments may also be a factor in why teachers focus almost exclusively on obtaining a reading level rather than on other, possibly more important, aspects of the assessment. In particular, this quote again demonstrates the overwhelming focus on students' reading levels. According to Marisol, it is "a waste of time to assess" when she knows a student cannot master the next reading level, thus viewing the running record as a summative assessment. This stands in contrast to the stance that this assessment is an opportunity to gather more information about students' developing reading skills. In this sense, the way in which running records are administered and analyzed parallels the ritualized assessment practices which frame the use of the ELA data for emergent bilinguals. The twin pressures of limited time and fulfilling school mandates in effect transform this formative assessment into a summative one. In other words, the latent potential of this formative assessment to impact teacher knowledge and instructional practices for emergent bilinguals is lost. Rather, teachers are effectively using this formative assessment to report results, thus making this formative assessment a summative one. When teachers and school administrators are focused almost exclusively on reading levels rather than the analysis of reading behaviors (through miscue analysis), then teachers are not using this assessment formatively.

In fact, many teachers were aware that running records were not being used to their potential and therefore were not contributing to teachers' knowledge of students. Kai surmises,

“I’m not letting the running records impact my teaching the way that I think it’s meant to be.” (School #2, Grade 4, 11/14/10). Few teachers learn how to analyze running record data in general, and consequently fewer, if any, have the knowledge and skills to adapt their use of running records for the benefit of emergent bilingual students’ learning needs. For these reasons, the process of analyzing emergent bilingual students’ reading behaviors and miscues becomes a puzzle, and therefore, it is easier for teachers to focus on students’ reading level. Kai describes her lack of knowledge about how to analyze running records and the subsequent effect it has on developing instructional practices from this knowledge:

I haven’t been proactive in trying to move a student along. Mostly I just – we just keep on reading together and I tell them to keep reading. Part of that is true, the kid will keep reading and they’ll just get better, but I don’t really have many tricks up my sleeve, to really pull those readers along and really advance them. I’ve heard other more experienced teachers really talk about, “I pulled him for more guided reading and I’m working with him really hard to get him up to the next level.” There seems to be something really intentional there and I just don’t know what they are doing. (Kai, School #2, Grade 4, 11/14/10)

Kai’s story is important because it is symbolic of the struggle that teachers have in using data, regardless of whether their students are emergent bilinguals or not. Her narrative points out that teachers need to know how to link teacher knowledge to instructional practices that match students’ instructional needs. In fact, one of the respondents from the survey wrote the following, “with enough time and development these assessments [running records] could provide teachers with an overview of student development that goes beyond increasing levels – adding a whole new depth to the data on each student.” Using data to shape instructional practices is not automatic; rather it requires both time and guidance. Susanna, a former 5<sup>th</sup> grade teacher and now coach at School #3, remarks on her journey to using data effectively. She says:

I wasn't immediately able to use data. I'd say it took me a couple of years to become efficient enough at gathering data to actually have the time to put effort into using it. It takes a while. (Susanna, School #3, Coach, 11/26/11)

Aside from the complex processes required to use running record data and transform the data into useable knowledge, there are other factors which cap teachers' use and acquisition of knowledge from running records. An additional barrier in three out of the four schools where teachers were interviewed was the availability and quality of assessment materials. Some teachers expressed that the administration of running records was fraught with serious problems. At one school, complete "sets" of running records were not available in either English or Spanish. In other words, these kits missed some of the sequenced texts that students need to read in order to be assessed. The following teacher describes how running record assessments were completed by piecing together parts of assessment materials from different assessment kits:

We actually are using a couple [different reading record kits], because we needed to get the assessments done and we didn't have a complete set. There's a variety. (Emeline, School #1, Grade 3, 11/7/11)

Furthermore, the same teacher maintains:

The diagnostic was only successful in getting us off the ground because we needed a starting point. My co-teacher and I aren't entirely satisfied that those levels that we gave them are their level, especially since we didn't use the same assessment for all of them. (Emeline, School #1, Grade 3, 11/7/10)

These two quotes taken together demonstrate that although running records have the potential to collect more frequent and specific information about emergent bilingual student reading, this capacity is often unrealized not only because of the lack of teacher knowledge about how to use them, but also because of more mundane issues such as the organization of materials. In these two specific quotes, Emeline points to the problem of improperly and inconsistently

using materials from various running record kits. This implementation practice then renders the assessments invalid and of limited use because texts from different kits may lead to different results.

Teachers identified a range of factors which impacted the extent of student knowledge they could glean from running records. Time, mandates, the need for guidance and modeling on how to move from assessment to instruction, as well as the availability of materials limited how much teachers analyzed the running records and hence their resulting levels of student knowledge.

Additionally, although teachers reported learning about students through running records, this knowledge was not specific to how these students' reading was deeply influenced by their bilingualism. Teachers focused on their students as readers in either English or Spanish, hence considering their students as two monolinguals rather than a composite bilingual person (García, 2009). As such, the knowledge that teachers acquired can be considered generic rather than bilingual specific.

In fact, gaining knowledge about emergent bilinguals' biliteracy was not a concern of teachers. The missed opportunity to learn about emergent bilingual students' reading can be attributed to, in part, the fact that teachers are under acute pressure from school administrations to administer and report data about the students' English reading. Because of these pressures, teachers placed an unbalanced emphasis on English reading levels. Lastly, the ability to conduct a thorough analysis of the running record data requires both time and guidance, both of which are not provided to teachers. The next section further examines the acquisition of *assessment knowledge* from running records.

## **Owning the Assessment: The Acquisition of Assessment Knowledge**

The previous section on student knowledge, unfortunately, sets the stage for the remaining types of knowledge studied through this research. The way in which the running record process is being implemented is far from its formative intent. As such, a pattern emerges. Teachers are not acquiring substantial levels of any type of knowledge from the analysis of running record data, especially with regard to understanding student bilingualism. As such, the relationship between assessment data of emergent bilingual students and assessment knowledge is very weak.

Assessment knowledge is defined as an understanding of the ways assessment use can lead to practices which support students (Heritage, 2007). Assessment knowledge is critical in the assessment-instructional cycle as it facilitates teachers' understandings about how to administer and analyze assessments in the service of student learning. Manifestations of assessment knowledge can be seen in the modification or adaptation of an assessment with the purpose of more accurately capturing student knowledge. It can also be demonstrated by analyzing assessment data with an eye for language development.

Most teachers interviewed did not do either. However, some teachers were cognizant of the gap in their assessment knowledge and called for professional development opportunities to enrich their understanding of how to make assessment more useful in understanding students and in shaping instruction. Suzie, a fifth grade teacher, asks for more guidance, through modeling, about how to make the link between assessment data and instruction. She says:

I guess I would like a little bit more training. What does a really strong reading conference look like? Have that modeled or, "How do you sit down with your running records and break down the data in a meaningful way?" That's not something that even

though it was explained to me how to do it, and how to interpret it, I feel that it's rare that I'll sit down with a running records and make a list of recommendations for student or somehow use that data in a reading lesson. I don't feel like the running records have a particularly strong connection to the mini-lessons that I choose to do. (Suzie, School #2, Grade 4, 11/12/10)

This quote highlights that teachers need the skills and practice to understand how to use assessment data to shape instruction, even before analyzing this data according to the needs of bilingual learners. It hints that teachers are lacking the basic skills to make data valuable instructionally for any type of student. Furthermore, although teachers admit that one of the functions of assessment is to see progress, because of the undue emphasis on reading levels, teachers may only recognize progress through change in reading levels rather than in more detailed understanding of student reading obtained through miscue analysis. Consequently, since teachers are not involved in the in-depth analysis of student reading, teachers are not aware of how students have progressed or even how to enable further progress. The finding of teachers' deficit of assessment knowledge, while sobering, clearly indicates a direction in which schools can focus their professional development and improve the use of assessment by teachers.

Although most teachers did not indicate acquiring assessment knowledge from running record data, a sub-set of teachers across the schools described ways in which they used assessment knowledge they already possessed in order to understand emergent bilinguals' reading development. These teachers were present at all of the schools and had taught for more than five years. They viewed the running records as malleable and open to adjustment. It is important to note that the teachers who did hold this viewpoint were those who already possessed a moderate level of assessment knowledge. Furthermore, they blended assessment knowledge with their knowledge of bilingualism in order to fashion accommodations of running

records. This point is important to this study because it highlights the interconnected nature of teacher knowledge. The following narratives also signal that teachers need to already possess some level of knowledge in order to propel their acquisition of other types of knowledge further, underscoring the importance of guidance and support in the use and analysis of data of emergent bilinguals.

Victoria, a fourth grade teacher at School #3, states that she views the running record assessment as open to adjustment. These adjustments, although not detailed in this quote, are based on students' level of receptive language. She may change the types of questions she asks of her emergent bilingual students, ensuring that they can answer them so that she can get an accurate measure of student comprehension. She describes her attitude towards this process:

I think that we have a lot of authority in how we administer the assessments and I mean even in terms of rephrasing the questions or asking different questions that I think will tell us more or based on whatever they read. I don't feel scared to do that. I feel very comfortable doing that. (Victoria, School #3, Grade 4, 5/17/11).

Another teacher at another school, Aura, speaks about her ability to rework running records as well. This quote highlights the importance of this work in the company of other teachers in order to build off the knowledge that other teachers possess as well as to develop cohesive assessment practices at the school. She states:

We are constantly tweaking the actual assessment, reflecting on whether it's working or not – what's working, what's not working, so we can stay as current as our students are and we can grow along with them to help them the best we can. (Aura, School # 2, Grade 3, 11/12/10)

These two quotations illustrate that running record assessments are flexible in their administration. Teachers can make adaptations that suit student's language abilities. Therefore,

these informal assessments can be modified to more accurately account for emergent bilingual reading. However, teachers are frequently at a loss as to how to make these determinations.

Another potential way in which teachers may make adaptations to the running record process is through the analysis of miscues. Students who are reading in their second language may make errors that have less to do with their reading abilities and more to do with their language development. Elisa speaks about these errors and how they are usually not accounted for when conducting the miscue analysis:

They [the assessments] really don't take into account a student's level of language acquisition, so a lot of the mistakes they are making, well, ya, they should be making that mistake because they are only, this is only their first or second year, speaking a second language. Those are mistakes that you would expect from that level but they really don't take those mistakes into account because they are geared to monolingual speakers of a language. (Elisa, School #3, Grade 4, 5/17/11)

Elisa's quote speaks to the lack of discussion among teachers in dual language bilingual programs about how errors that emergent bilingual students make may be the result of reading in an additional language. Not taking these into account and counting them as errors will lead to students being assessed at a lower level than they may actually be.

The following teacher, Amy, a fourth grade teacher at School #1, speaks more in depth of the accommodations that are necessary when examining student miscues when reading in a second language. The caps in the text denote a brand of running record kit.

I definitely make accommodations with XXXX because our ELLs have certain ways of saying certain words. I definitely note them, but depending on the miscue I may not actually count it. If the kid says, "walk-ed," I would mark that and I probably would also count it as a miscue. But if a kid were to mispronounce a word during reading that he mispronounces when he even says it in oral conversation, then I won't mark it as an actual miscue. I'll note it, but I won't mark it as an actual miscue. So that's something

that I do to accommodate English Language Learners. (Amy, School #1, Grade 4, 11/26/10)

Amy's quote is more specific, but demonstrates confusion over what actually is a miscue for an emergent bilingual who is simultaneously learning English and learning to read in English. In her example of "walk-ed" it is important to determine whether the Spanish-inflected pronunciation of the word interferes with meaning. For some students it may, and for others, it may not. However, what this quote points out is the pressing need for teachers in dual language bilingual schools to talk about these issues in a way that takes into account language development and to develop cohesive school-wide practices. It is these types of discussions that would simultaneously enrich teachers' store of assessment and bilingual knowledge.

In the next section, I will turn to bilingual knowledge. Teachers spoke at length about the issues regarding bilingual assessment. These issues provide insight into the gaps in bilingual knowledge that ultimately limit teachers' understanding about how these assessments can be best used for emergent bilinguals.

### **Seeing the Whole Child as a Reader: The Acquisition of Bilingual Knowledge**

Bilingual knowledge as defined in the literature review as the knowledge of understanding how two languages develop simultaneously and how to facilitate the development of language and content in two languages. Bilingual knowledge was absent from most teachers' narratives of running record use. Most often the question of how teachers used running record data to learn about emergent bilingual students' reading development was met with either a negative response or with the response that these two sets of data were examined separately and never merged. Despite this, teachers had a tremendous amount to say about their beliefs about

language and how learning in two languages impacts assessments as well as their teaching. This section of this chapter details both barriers and teachers' beliefs about students' language learning.

The following two excerpts give some examples into the sorts of responses teachers gave to the question — How is running reading data examined bilingually? The first quote confirms that teachers often have the singular focus on obtaining a reading level from this assessment. This exclusive focus on reading levels distracts from the potential of using running records to be used bilingually, and hence yield bilingual knowledge.

Laura: Do you ever look at your student data or assessments bilingually?

Jaime: What do you mean?

Laura: You will take one student's data and look at the running record in English and then in Spanish and then think about are they using the same strategies in both languages.

Jaime: Not in terms of the errors they've made or what skills they do have, just in terms of that letter, that level. That's pretty much it.

(Jaime, School#1, Grade 3, 11/26/11)

In the second excerpt, Elisa's response demonstrates that the analysis of running record data remains separate for English and Spanish data. Therefore, the teacher does not acquire a complete picture of the student's reading in both languages. Rather, as she states, her knowledge of student's reading in English and Spanish remain "compartmentalized."

Laura: So how do you use, if you have a reading record in English and a reading record in Spanish, how do you use it to understand the students in terms of their biliteracy?

Elisa: That's interesting. I feel like I compartmentalize them. I think of their English and I think of their Spanish because there are nuances in the language that I don't like to overlap sometimes. (Elisa, School # 3, Grade 4, 5/17/11)

Elisa's response is unique because of her claim, although unspecific, that there are nuances in the language that she wants to preserve. However, it also is representative of how

most teachers view running record data in English and Spanish as two separate entities. Furthermore, they are not accustomed nor encouraged to merge them in their analysis of emergent bilingual reading. Both of these responses demonstrate that the current way in which teachers conceptualize and analyze running record data does not take into account students' bilingual reading development. It is important to note that when speaking about running record analysis, I refer to miscue analysis, the heart of this formative assessment that helps teachers to understand student reading skills and thinking contextually.

The following teacher's story summarizes how current practices do not facilitate the comparison of student reading abilities across languages nor encourages a unified understanding of student reading. She states:

For the most part it's their English development and their Spanish development, although I think it's a great idea to look at them side by side, but I wouldn't know on demand that X student has the same issue in both languages, although the information could be there in the running record. But I do have a general idea of how students are in their language development in both languages. I know who the ELLs are I know who the SSLs are and I also have a chart that has their level in English and Spanish next to each other, so I can see really easily how their levels compare, but that doesn't have the specific data of what issues they are having in reading in Spanish or what strategies are they using in Spanish versus in English. (Suzie, School# 2, Grade 5, 11/12/11)

Despite the widespread trend among teachers to not view their data bilingually, a few of the teachers described the ways in which they did. The teachers who did engage in a bilingual analysis of their students emerged with a greater sense of their students' biliteracy and with the ability to provide cohesive instruction in both languages. The following four quotes are presented not to demonstrate an overall trend, but rather to provide a glimpse into how a bilingual analysis of running record data may look.

When asked if she examined data bilingually, only one teacher, Beatriz, responded with a resounding, “yes,” that it is important to look at students bilingually. Beatriz remarked:

Yes, absolutely, because if I see that their abilities in one language usually there should be a somewhat balance, similarity. So if it's an English dominant student learning Spanish, there shouldn't be a huge discrepancy between the levels. Obviously, if he or she is learning the Spanish, it should be slightly below what they are able to do, but I think if they are able to transfer all of the skills that they are learning in one language to the other, it should be ... so yes, absolutely I need to look at my students in both languages, think of them in what they are doing in English and what they are doing in Spanish, because it's the only way that I can help them. (Beatriz, School #1, Grade 4, 11/26/10)

Beatriz talks both about comparing student reading levels in English and in Spanish, but also goes into understanding the reading skills which students possess in both languages. This closer look at emergent bilingual student reading provides teachers with a more accurate understanding of their students' developing bilingual reading abilities and therefore provides a solid basis for instructional practices to aid *bilingual* reading development.

In the following quote, Judy, a fifth grade teacher at School #4, describes how looking at data in both languages allows her to determine the factors that account for her students' progress or lack of it. Judy states:

Well the good thing about doing running records in English and in Spanish is that we get a clear indication if a child is biliterate, which is great for us to know because we can eliminate if there is a learning disability, not eliminate, but we can judge if it's a language problem or it's a disability. We are taking their language acquisition in both languages into account when we are doing our assessment. (Judy, School #4, Grade 5, 12/1/10)

Another way in which bilingual analysis affords an important perspective about student learning is by informing how to support student learning in both languages. Zoila speaks about how she does not compartmentalize each student into a reader in English and a reader in Spanish.

Rather she considers the unified reading development of a given student in English and in Spanish. She states:

Actually I look at the whole child. I have a couple who are really fluent in Spanish more than English so I make sure they are getting the same praises. If I see that they are doing something right in their native language, I am not going to dismiss that just because it's English week. I am going to make sure they know, "you are doing it well in Spanish, awesome, I like the way that you use adjectives. That's a nice adjective to use." (Zoila, School #2, Grade 3, 11/12/10)

Zoila's consolidated vision of a student who is developing in reading provides her with student knowledge and bilingual knowledge in order to base her instructional practices. The following quote features Zoila describing how she merges the student, assessment and bilingual knowledge after analyzing running record data in two languages. In the following quote she describes how she works with children during guided reading based upon running record analysis.

One thing that we do is have conversations about the book before. Like give them a book introduction and then we start talking about the book. They have to share information that they learn about the book, how it connects to their lives, the connections that they may have with the book. And as they do that, you are able to learn from them as they are telling you verbally. Wait a minute, are they using the right verb tense? Conjugation? Are they using the right articles? Are they using the right adjectives? Transitional words? And so forth. So if I notice that they are not using a lot of transitional words, for example, what I do is that I check the writing to see if they are there. But if I see that they have been exposed to it and they are not using these things or these words or whatever I am looking at then one thing I started to do ... is that I noticed that they are using the word, "then, then, then, then, then." So we did a lesson briefly on transitional words. You don't have to say, "then, then, then, then" or "and, and, and" all the time. (Zoila, School #2, Grade 3, 11/12/10)

Through Zoila's narrative, we see the entire assessment-instructional cycle modified for emergent bilingual students. She moves from the acquisition of specific knowledge about emergent bilinguals to fashioning instructional practices attuned to their needs works.

Unfortunately, this pathway was not the norm for the majority of teachers. Rather, the majority of teachers were locked into a process of collecting and reporting data that rendered running record data to be both generic and summative in nature, as described earlier.

If there is such great potential and positive regard for running records to undergird bilingual reading instructional practices, why is this road not taken more often by the teachers interviewed in the sample? Answers to this question, can be partially found in the critiques of the time and reporting pressures featured in the section about student knowledge. However, these explanations are not enough. Further reasons can be found in understanding the differences between English and Spanish running record assessment practices. The following section details the ways that teachers identified Spanish running records as differing in quality from English running records.

### **Assessing Students through Spanish Running Records: Inequities from the Start**

One critical finding that emerged from the interviews is the inequity between English and Spanish running records assessments. Teachers from all four schools stated that the quality of Spanish reading assessments did not match what was available to them in English. The following teacher attests to the difference in quality between English and Spanish assessments as well as in different running record materials across products. The caps embedded in the text are intended to denote a particular brand of running record kit.

I think the other issue that I run into is that it seems that the quality of the materials for assessment is much higher in English than it is in Spanish. So we used to use the XXX in English and Spanish and we've done away with that because YYY is just a higher quality sort of product. (Ariana, School #2, Grade 4, 5/17/11)

At another school, one teacher has designed the entire Spanish reading assessment for all dual language bilingual teachers to use. While this is impressive, teachers noted that the assessment created by the teacher is not on par with the English reading assessment. One teacher at the school states:

In Spanish, I use the books that were carefully selected by Ms. Torres, the third grade teacher. She put a lot of effort into doing all the running records. She did a great job, you know. The only problem is that we only had one book per each level so if the student doesn't master it, and sometimes it's the same book for two levels, so sometimes if the student doesn't master that level he gets to read the same book the next time, he already knows the story. (Marisol, School #4, Grade 4, 12/1/10)

In this case, because there is only one book per level, students who are re-tested in Spanish at a specific level, are assessed through reading a book that they have read previously. This may compromise validity of the running record assessment, resulting in an inaccurate evaluation of student's Spanish reading abilities. These quotes suggest that the results in English and Spanish that teachers collect from their students may not be comparable. This finding can be generalized across the four schools where teachers were interviewed, as teachers at all of the schools had different running record kits in English and Spanish or, as in one case, the Spanish kit was created by a staff member. Therefore, it can be said that all teachers at these dual language bilingual schools worked with very different running record materials in English and in Spanish.

The issue of equity between running record materials in English and Spanish does not end in the selection of quality assessment materials in both languages. In addition, it is important that Spanish reading materials do not favor one Spanish speaking population over the other. For example, in the following quote, one teacher noted that the texts provided within a given running

record kit may favor Spanish vocabulary hailing from one country over another, thus advantaging one Spanish speaking population over another. This teacher remarks:

It's so hard because with Spanish there are so many different types of Spanish that is spoken. And so testing a number of students who come more or less at a level K and the story that they have to read is *Hielos de fresa* [Strawberry Ices] and these kids are like *charola* [tray]? They don't know what that means. They don't know what *rabito de fresa* [strawberry slice] is and they call it something different. I think it's really hard with so much Spanish that is spoken, if a kid comes from the DR [Dominican Republic] and he calls something one thing and a kid comes from Mexico and calls it another thing, they are both correct but how do you assess that? Like the kid from Mexico is lucky because the assessment happens to have words that he knows. (Emeline, School #1, 11/7/10)

As the teacher above commented, because the texts in running records kits in Spanish may contain solely vocabulary from one regional variety of Spanish, the running record assessment may not accurately capture students' reading ability – even for students whose home language is Spanish. One of the main issues with having running record kits of different quality in Spanish and English is the quality of information garnered through these different assessments. For example, teachers may gain more information from English running records than Spanish ones because the English running record assessments are better constructed. In turn, teachers may trust these more and, as a consequence, be more likely use them for their own learning and to shape instruction practices.

Another finding stemming from the lack in equity in assessments is many teachers' confusion about how to move from Spanish running records to Spanish reading instruction. Teachers repeatedly stated that Spanish libraries and resources do not match the needs of their students. Therefore, although they may have a sense of their students' reading abilities in Spanish, they do not have the resources to support students' Spanish reading. The following teacher attests to this inequity in materials and its effect on instruction:

I do feel like the growth is a little bit different [between English and Spanish] because if you look at my libraries I do have fewer books in Spanish. And we are doing a particular genre in writing or social studies, it's twice as hard to be able to get a book in Spanish and also translations are sometimes terrible. (Aura, School #1, Grade 3, 11/12/10)

While assessment materials in Spanish may be of a lower quality than those in English, even when teachers have assessed students in Spanish, they do not have the materials needed to instruct them in Spanish. These two issues add to the reasons why teachers may not analyze running record data in English and Spanish bilingually – because these two data sets cannot be compared.

For a number of factors, dual language bilingual teachers do not examine, through miscue analysis, running record data for emergent bilinguals *bilingually*. As previously mentioned miscue analysis is an essential step in understanding how students' errors indicate what skills and strategies students employ while they are reading. Not engaging in miscue analysis further contributes to the lack of bilingual knowledge that teachers possess. As described in previous sections, what became apparent from teachers' narratives is that the development of teacher knowledge is mutually dependent; all types of teacher knowledge in this study are interconnected and are necessary in order for teachers to develop new knowledge.

In the following section, findings are shared about what teachers said about language and literacy transfer, as well as what teachers' conceptualization about this topic may reveal about teachers' bilingual knowledge. Many teachers referred to how students could transfer their strong reading skills from one language to develop reading skills in the other. These teacher narratives provide an insider glimpse into how teachers instruct with the goal of bilingualism in mind.

## **Language Transfer: Bilingualism Takes Care of Itself**

Many teachers interviewed evoked the power of transfer between languages. In fact, one of the pillars of dual language bilingual programs is that students' academic language development will be enriched if students are able to transfer the language-knowledge that they have between one language and the other. One teacher states:

I do expect them, whatever skills we teach them, and in whatever language, to be transferred to the other languages as well, so that the skills should be the skills. (Milena, School #4, Grade 3, 12/1/10)

Teachers repeatedly mentioned the same belief voiced by Milena – that language skills will automatically transfer from one language to another. Elisa speaks about this fluidity between languages. She states, “I generally feel that if you’ve had a certain amount of exposure to each language and you have strong skills in one, then they’ll transfer to the other” (School #3, Grade 4, 5/17/11).

Nonetheless, absent from the interviews were specifics about how this transfer would occur within the classroom setting or how knowledge from assessments would aid teachers in facilitating transfer. Teachers, rather, remarked that the transfer would happen automatically, above all, when the student conditions were “right.” This includes, but is not exclusive to, students who come from a home where Standard English is used or students that have high literacy levels in their home language. Victoria, another teacher at School #3, expresses the following:

I have a student who is a non-native Spanish speaker and he’s so bright and this is where you see the transfer. His reading level in English is very high – it’s at grade level or higher. His Spanish is like second grade. But because of the way he just processes language and thinks, even though he cannot decode all those words, his ideas are solid

and he helps all of those kids who in the same Spanish level, who maybe speak Spanish more fluently than he does, but they cannot think and comprehend the deeper meaning of the text that he can. He brings a great, great perspective to that group. (Victoria, School #3, Grade 4, 5/17/11)

This remark is not exclusive to this teacher. The sentiment among many teachers is that reading skills will transfer from one language to the other when student skills are high. In a way, teachers believe that bilingualism develops independently or “takes care of itself,” instead of resulting from teaching practices geared to enhance transfer between languages.

This reliance on transfer to happen by itself is ingrained in many teachers. Zoila, a teacher at School #2, says, “there are elements in one language that will definitely not be in the other language because one language is stronger at this point. So I just let it develop naturally and through other kinds of support hopefully students will get the language support” (School #2, Grade 3, 11/12/10). These quotes taken together point to the finding that teachers rely on the belief that students will transfer reading skills between languages without teacher instruction. However, the transfer of reading skills between languages does not occur on its own (Proctor et al, 2010; Baker, 2011). Perhaps this lack of engagement in understanding how transfer does contribute to bilingual reading development can be attributed to teachers not using miscue analysis to examine emergent bilingual students’ running record results. Since they do not practice miscue analysis, nor miscue analysis adapted for *bilingualism*, teachers are not building their knowledge of student bilingualism or of student bilingual reading development. Rather, teachers may view bilingual knowledge as acquired through study in professional programs rather than built in-situ alongside students.

The next section differs from the ones above in that it highlights one teacher's responses about pedagogical content knowledge of reading, and especially of biliteracy. Her insights into knowledge of biliteracy are important to the discussion of bilingual reading development because it provides a backdrop for seeing how the teaching of reading, when combined with bilingual knowledge, can differ from what is currently in place at many dual language bilingual schools.

### **Pedagogical Content Knowledge of Reading Bilingually: A Singular Voice**

In this dissertation, pedagogical content knowledge is defined as the knowledge required for teachers to instruct students in a specific *domain* – in this case reading. Throughout this thesis I have claimed that the teaching of reading and analysis of reading data for emergent bilingual students differs from that of English-speaking monolingual students. Although most teachers did not speak about pedagogical content knowledge at all, one did begin to speak about how reading in English and Spanish compare. This teacher's quotes will be featured, not because they are representative of the sample interviewed, but because they provide insight into possible ways pedagogical content knowledge can be conceptualized for teachers of emergent bilinguals. Beth, a third grade teacher, begins by stating that bilingualism is not at the forefront of educators' minds. Rather, teachers think of teaching as divided; they plan for teaching content in English and teaching content in Spanish.

I don't think in this school we do very much with testing language ability. I think we also don't do much of teaching language. We teach in Spanish, we teach in English, but we don't necessarily teach the language. (Beth, School #1, Grade 3, 11/4/10)

The point made by Beth goes beyond teaching in either English or in Spanish. She highlights the importance of understanding students' development in academic areas, such as reading, alongside the critical role that language plays in student learning. Furthermore, her

comment points to the unique nature of teaching in a dual language bilingual school. In these schools, teachers are teaching content and language. However, often teachers are engaged in teaching content through an additional language and leave out the explicit development of the additional language. This leads to the misconception that the teaching of any content area (such as reading) is the same regardless of the language. In the following quote, Beth breaks down this assumption. Beth comments on how the reading process differs between English and Spanish.

She states:

What exactly is reading in your second language? Is it about correct pronunciation? Is it about getting meaning? So I think that some people expect the child to read out loud phonetically in the second language, where I believe more in the meaning. And I don't know that I'm right, but we've had kids that are Spanish speakers who in English it just comes out as garble and you think that this kid understands nothing, but then later you ask them what happened and they can explain it. (Beth, School #1, Grade 3, 11/4/10)

This quote provides a very promising starting point for thinking about the assessment of emergent bilinguals and how instructional practices can emerge from this type of bilingual-specific pedagogical content knowledge. In the following quote, Beth specifies the differences between reading in English and Spanish and the implications these differences may have in emergent bilinguals' reading development. Beth asserts:

English is very, very predictable based on what's happening in the story. With the picture and that first letter you can basically predict the word at the very low levels. Whereas in Spanish the ending of the word is so important, you know what the verb tense is ... is it a male noun, female noun that I think particularly in the beginning levels reading in Spanish is much harder. I mean I worked for years as a reading specialist and I found that Spanish reading was much harder even though you can phonetically sound it out. (Beth, School #1, Grade 3, 11/4/10)

Beth challenges the commonly-held idea, that Spanish reading is “easier” than English reading. The idea that Spanish reading is easier to learn than English reading is frequently held by many bilingual teachers. This assumed ease in learning to read in Spanish is attributed to the fact that the Spanish language is phonetic and therefore, more readily decodable for students. In this sense, the term, “decoding,” is used interchangeably for reading. In spite of this, according to Beth, other elements of the Spanish language may make it harder for readers in Spanish. If a language is phonetic, readers may primarily use visual and meaning cueing systems to read. Given that, according to this teacher, since Spanish may be syntactically more complex, readers would have to use their syntactical cueing system as well to also read. Beth continues by questioning how to assess readers who are developing reading skills in language they do not speak at home and in which they are not proficient:

You know if you want to get meaning from it [Spanish reading]; I think it's much harder. The words are a lot longer and so then again, my question is on a running record, “Do we count the kid wrong, who got the wrong the verb ending, but got the proper verb? Like if they read *habló* instead of *hablaba*? Yes, it is incorrect, but if they get the meaning? But they didn't read to the very end of that word. In English you don't have to read to the very end of the word. Some teachers in our school count that as wrong. I think for some teachers it's very much black and white. It's either right or it's wrong. I'm much more into shades of gray. And also I look at assessments as hopefully telling me what I need to teach them, not telling me what they don't know. It's not the end. (Beth, School #1, Grade 3, 11/4/10)

General miscue analysis assumes that students read in the home language, for which they have proficiency. Beth speaks about the issue that all students in a dual language bilingual program have — that they are learning to read in both their home language and in a second one. As such, miscues will occur. However, as Beth cautions, miscues must be interpreted with an eye for language. Used in this way, they have the potential to reveal as Beth says, “what I need

to teach them.” This quote demonstrates a promising direction for dual language bilingual teachers in the use of running records to understand students’ intertwined development in reading and language.

In addition, the quote above attests the importance of establishing consistency among teachers in the ways in which they assess students in their second language. Another point made by Beth’s narrative is that although periodic running records are formalized, meaning that teachers administer them according to instructions from a kit, they are not administered the same way among all teachers. In other words, the accommodations that teachers may make for language learners are made on a teacher-to-teacher basis and are far from being implemented regularly and purposefully.

Beth’s quotes taken together show the flow between the various types of teacher knowledge studied in this research. Her understanding of bilingual reading development shows how she came to a complex conceptualization of reading in English and Spanish. This knowledge, in turn, provides the foundation for assessment practices that take into account student bilingualism.

Even though most teachers did not build a similarly multifaceted understanding of student bilingual reading development, this type of teacher development of integrated knowledge is possible. However, this would require that teachers step out of ritualized assessment practices and work intensely on understanding how to work with data starting from student needs and language development. In the next section I move on to describe the instructional practices that teachers attributed to the analysis of running records.

## **Instructional Practices: Assessment Behaviors Disconnected from Data**

Teachers spoke about instructional practices resulting from analyzing the assessment data from running records. Four main instructional practices were frequently mentioned by teachers. These were: 1) grouping, 2) the design of mini-lessons, 3) vocabulary, and 4) student feedback. Each of these will be discussed separately.

### ***Grouping***

As was detailed in the previous sections, teachers primarily focused on student reading levels obtained from running record results over miscue analysis. They were required to report these to school administration periodically. As such, they were used by all teachers. Suzie, a fifth grade teacher speaks about how she uses the student reading levels in order to group students. In addition, she speaks about the types of instruction within each of these groups that may attend to student needs. She states:

Well, we might use the running records to make guided reading groups or literature circles. So by grouping the students by their level then you can target issues specific to that level of development, so if they are reading below level, then maybe providing more support and more ways of decoding difficult words. Whereas, if they are reading fluently at above grade level then they might be able to take on more challenging questions – more inferencing questions. (Suzie, School #2, Grade 4, 11/12/10)

Although initially benign, grouping according to running record levels may not be best suited for emergent bilinguals, as it has been shown to be the case for proficiency level based on summative assessment data. For example, in the case above, if emergent bilinguals are deemed to be reading below level, they would be with a group of students who are being taught to decode, regardless if their reading level was due to a weakness in decoding or in language ability. This is particularly a serious concern if students are being assessed by teachers without

having a clear and consistent policy of which miscues are counted as such, and which are attributed to learning in a second language. Unless teachers analyze how language impacts students' miscues, the danger of emergent bilinguals being placed in groups that do not meet their instructional needs is present.

The following quote demonstrates how teachers may implement instructional practices within group instruction which are derived from a set of fixed practices intended to move student to the next level. This seems to reflect what teachers do with summative assessments such as the ELA, as described earlier. Sameer, a fourth grade teacher describes his view of group work based on student levels:

We have students who are grouped at the same level and like I said, we have bands [of instructional practices] at each level. When we do group strategy lessons, we group all the N readers and we focus on one band. So you work on that band and hopefully when you do the next running record, they've mastered the N level. (Sameer, School #4, Grade 4, 12/1/10)

However, not all teachers use student levels from running records exclusively to group their students. Rather some teachers, like Victoria, use them initially, and then rely on observations to further group her students:

I think we use it as a starting off point. How can we group these students but then especially after we get to know the students after a few weeks or at this point after a few months, we are like, "OK, but we know this kid really needs to be in this group because those kids will support him more or because this group is always with a teacher." So we use it as a starting off point but it doesn't determine everything. (Victoria, School #3, Grade 4, 5/17/11)

This last quote demonstrates that although many teachers use student level as benchmarks to kick off their grouping, these groupings for some teachers may not be long-lasting. This reflects the limited degree to which teachers process and use data about their students from running records.

In all quotes, however, grouping as an instructional practice does not emerge from knowledge teachers acquire about their students from actual running records, but rather from a school-based practice that is required, a consequence of the data-based accountability culture. Although grouping is an important instructional practice, the exclusive focus on reading levels demonstrates that teachers miss the opportunity to learn about their students in more meaningful ways. Therefore, the link between the running record assessment data and grouping as an instructional practice does not reflect knowledge that teachers acquired from the assessment. Instead, the practice shows that teachers implement instructional practices based on set “ritualized assessment practices” much like those implemented around ELA data and instructional practices.

### **Mini-Lesson Design**

Teachers also spoke frequently about mini-lessons emerging from running record data. Mini-lessons are short lessons created to teach students before moving onto their own independent work. They are labeled “mini” because they are supposed to provide a quick demonstration to students, rather than a lengthy lecture. Although frequently mentioned by teachers as a way that they use the running record data, as in other findings, they were used generically without an eye for the needs of emergent bilingual students. Amy, a fourth grade teacher at School #1 states:

I look at my running records and I say OK, my kids are not for instance looking for a word inside a word. And that would really help them. They are not looking at the word construction and saying, ‘Oh, I know that and that can help me figure out the word.’ So being able to take that information [from the running record results] and do a mini-lesson the next day. (Amy, School #1, Grade 4, 11/26/10)

Amy describes a solid and valid practice for the teaching of reading. She notes that students are not finding word parts inside larger words in order to sound them out (for example, “out” can be used to read the word “pouting”). Therefore, she takes this information and teaches them this strategy in a mini-lesson. However, teachers did not use their knowledge of student bilingualism to fashion mini-lessons. Rather their design of these reflected a monolingual view of reading. Their mini-lessons were based on students learning to read mostly in English, and sometimes in Spanish, but not to become competent biliterates.

### **Vocabulary**

The introduction of vocabulary was another instructional strategy that teachers mentioned as guided by running records. Vocabulary was spoken about in two ways. On the one hand, teachers spoke about introducing vocabulary to students who are encountering new words in general. On the other hand, they spoke about the relationship of vocabulary knowledge to culture.

Teachers determined what vocabulary students lacked and provided it to them. This determination was loosely based on the running record. The assumption is that one of the greatest difficulties that emergent bilinguals will encounter in reading is vocabulary. Marisol, a fourth grade teacher, describes solid vocabulary teaching strategies, albeit generic.

When it comes to retelling we need to incorporate the vocabulary so we are trying to develop the language through groups. When they have the conversations about books, we go into the big group and everyone gets to express their ideas about the book and that’s according to their levels. (Marisol, School #4, Grade 4, 12/1/10)

From Marisol's response along with others like it, it is hard to determine whether the teaching of vocabulary is an instructional practice that emerges from running record data or rather reflects the belief that emergent bilinguals lack vocabulary and therefore need more work in this area.

Teachers also taught vocabulary in a manner that was linked to an understanding of language and culture. Although not all teachers taught vocabulary in this way, there was a small group of the sample that did. In the following quote, Suzie, a fifth grade teacher, remarks that she may adjust her teaching of vocabulary while she reads with children based on her knowledge of students' Spanish and country of origin. This quote, although not reflecting knowledge that comes out of running record analysis, demonstrates how teachers can fine-tune vocabulary instruction to meet the needs of emergent bilinguals.

So like in Spanish a lot of times it happens that the books are from many different countries. So the vocabulary can be very tricky, even if it's a level that they are capable of reading in general. So noticing if a group of students or one student or many students are getting tripped up on unfamiliar vocabulary words, even though it's [the book] the right level for them, kind of teaching into that moment and explaining how using the context can show them the word or kind of being able to differentiate between when I should stop my reading and lookup a word in the dictionary when I don't know it versus when is it OK to just skim over it and trust that it's not going to break apart my understanding of the book. (Suzie, School #2, Grade 5, 11/12/10)

Beth, a third grade teacher, also works with students on vocabulary while taking into account how vocabulary found in student reading may be embedded within cultural experiences. She describes how she previews the literature to ascertain concepts which may be foreign to students and discusses them with students before reading the book to them. In this example, Beth takes a concept, the *ratoncito*, and parallels that to the analog in the United States, the tooth

fairy. In doing this, she decreases the difficulty of certain culturally-specific concepts within the text, so that students can better access others:

If I know there's this book that first and second graders are going to read that refers to the tooth fairy, knowing that many kids who are native Spanish speakers might call it the *ratoncito* [little mouse], I'll go out of my way ... to bring it up, "What happens to you when your tooth falls out?" "Oh, the *ratoncito* comes." I'll say, "You are right. When I grew up we didn't have *ratoncitos* in my house, we had the tooth fairy." (Beth, School #1, Grade 3, 11/4/10)

Beth describes an instructional practice that adapts the reading experience for emergent bilingual students, so that they are challenged academically by the reading of the text, rather than by the interpretation of unfamiliar concepts. However, as with the other examples, although teachers did talk about the importance of vocabulary, they did not attribute learning about how to teach vocabulary to emergent bilinguals from the knowledge acquired from running records. In this first example, Marisol's vocabulary practices emerged from best practices that are performed by teachers regardless of student population. The last two examples point to a more sophisticated understanding of language and the emergent bilingual population. However, they are most likely practices that stem from bilingual knowledge acquired outside of the assessment experience.

### **Feedback**

Teachers talked about feedback as an instructional practice that they have adopted as a result of running records. However, when teachers mentioned feedback during the interviews, they spoke mostly about feedback that they provide students at times other than during or after administering a running record, such as during guided reading. Although guided reading is not a

focus of this study, it is mentioned here for the purpose of highlighting the ways in which teachers provide feedback to students.

Teachers in this sample wanted to move away from providing students with their running record level, a common-place practice. According to Elisa, providing students with their reading level does not improve their ability to make progress in reading. She states:

One of my pet peeves is that when people give students their reading level and their goal is to get to the next reading level, I feel that doesn't tell the student anything – like your goal is to figure out words in context and that's something really hard for you to figure out a word that you don't know from the text. So your goal is to really be able to figure out three words in context or goal is to make a really good prediction before you start based on what you know and then check your prediction afterwards. Or your goal is to then make a connection with the book. (Elisa, School #3, Grade 4, 5/17/11)

Elisa demonstrates how student feedback must be immediate and specific. These qualities of student feedback are important to acknowledge when considering how students can benefit from feedback of their running record data. Jaime says the following of the use of running record data and feedback:

I really just get the information for myself and then during the guided reading groups then I may give them that feedback: oh, you know what, go back and read it in a smooth voice or you have to read the whole word. I don't give them feedback immediately after the assessment. I would say something like, 'good job' or 'continue reading' because we just want the information. (Jaime, School #1, Grade 3, 11/26/10)

Jaime's quote reminds us that data from running records is being used summatively by teachers rather than formatively with students. Since many teachers view this assessment as a way to assess their teaching, and not for students, many teachers may not even consider how they use the data to inform students about their reading abilities and progress. Furthermore, as developed in previous sections, teachers are mostly focused on the reading level. However, the

reading level provides limited information to students about how to move their reading forward. For these reasons, running records are often limited in their ability to provide quality feedback to students, regardless if they are emergent bilingual students or not.

In sum, instructional practices were disconnected from data emerging from running records or teacher knowledge. Much like what was found for the ELA, teachers implemented practices that were either sponsored by the school, such as grouping, or implemented practices that did not take into account student bilingualism. Although teachers use the running records to inform instruction much more than they use the ELA, the use of running records to shape instruction is capped. These formative assessments hold enormous potential to provide teachers with detailed information about student bilingual reading and their biliteracy development. Since teachers do not engage in miscue analysis adapted for bilingualism, instructional practices do not target the development of bilingual readers. As such, these instructional practices do not hold much promise to impact these students' reading. Additionally, as one respondent to the survey writes, "a lot of the practices mentioned I would do anyway regardless of looking at assessments." This comment is important because it highlights that it is difficult even for teachers to attribute what instructional practice they implement to assessment data. Many of the instructional practices teachers currently employ would have been implemented regardless of assessment data.

### **Learning and Not Learning from Running Records: A Summary**

Due to their formative nature, running records are considered to better capture emergent bilingual reading development. Although dual language bilingual teachers hold these assessments in higher esteem than the ELA, they do not use them to their potential. In fact,

teachers acquire a very limited amount of knowledge that is specific to emergent bilingual student reading development. Data from the interviews indicate that although teachers acquire a fair amount of student knowledge from these assessments, this knowledge does not account for student bilingualism.

Teachers spoke at length about the contextual factors which limit the use of this assessment for the development of teacher knowledge. These include time pressure to complete the assessment in both languages for each student, the exclusive focus on obtaining and reporting student reading levels, and lack of guidance and support for how to analyze them. These contextual factors in effect render this formative assessment summative in nature because teachers view them primarily as a means of reporting student reading levels rather than as a means to assess their bilingual students' changing reading development. This finding must be underscored. Considering the very positive view that teachers hold of running records as a more accurate way to assess their emergent bilingual students, in practice running records are used as a *summative assessment*. Therefore, although highly regarded by teachers of emergent bilinguals, they are not being used to actively either contribute to teacher knowledge or shape instructional practices. Although the acquisition of teacher knowledge from this assessment was limited, teachers did identify instructional practices that emerged from this data. The following are findings with respect to teacher knowledge and instructional practices resulting from the analysis of running record data of emergent bilingual students:

1. Comparatively, teachers learned more from running records, a formative assessment, than the ELA, a summative assessment. Regardless, the knowledge that teachers acquired from running records was generic and reflected both a monolingual use of this assessment, as well as entrenched assessment practices.

2. Teachers report acquiring student knowledge above other types of knowledge. However the student knowledge that teachers acquired by teachers was not specific to the fact that these students were emergent bilinguals.
3. Teachers did not acquire substantial levels of knowledge about assessment, pedagogical content knowledge, or bilingualism, through the use and analysis of running record assessments.
4. A constellation of factors including school mandates, a singular focus on the student reading level, the time required to complete the assessments, and the need for guidance in order to aid teachers to analyze these assessments with respect to language, impacted teachers' ability to gain knowledge from these assessments.
5. The almost exclusive focus on obtaining and reporting a student reading level within the current high-stakes testing climate largely rendered these formative assessments *summative*.
6. Instructional practices that were implemented were disconnected from the knowledge that teachers acquired from the analysis of running record data of emergent bilinguals. Instead, instructional practices were shaped by teachers' prior store of knowledge and/or guidelines for best practices.

## **General Summary**

Although summative and formative assessments are very different in their format, administration, and the data they provide, they are used by teachers and school administration in very similar ways. Both of these assessments are used in ways in which they were not intended to be used. This is a result of teachers being caught within school-sponsored ritualized assessment practices that guide the analysis of each of these assessments. These ritualized assessment practices dictate the ways in which assessments are collected, how the data is analyzed, and how it used by teachers to shape instruction. Ultimately, these ritualized assessment practices are derived from test-based accountability measures, in which improvement

and success is defined as improvement on these very assessments. In this way, all assessments, formative as well as summative, are defined as valuable only when teachers can see discrete improvement in assessment results. Therefore, *teachers are expected to use summative assessment data formatively and are using formative assessment data summatively*. This mismatch between assessment purpose and use curtails the knowledge that teachers can gain from either and the potential for these assessments to exert a positive influence on emergent bilingual student learning.

Overwhelmingly teachers perceive large-scale standardized tests as unfair and incapable of yielding valuable information about emergent bilingual students' reading development; yet, these results are actively used by teachers to identify patterns of performance. Running records are largely perceived by teachers as the preferred way in which to evaluate the reading of emergent bilingual students. However, the ways in which data from running records are both collected and analyzed are not specific to the needs of emergent bilingual students. The full potential of running records is not used to accommodate or adapt to the characteristics of emergent bilingual reading because teachers are almost exclusively focused on obtaining and reporting reading levels in English and are being used summatively. Furthermore, findings suggest that teachers are operating on the idea that bilingualism takes care of itself – that students will become bilingual and biliterate simply because they are instructed in both languages and that teachers do not have to teach toward the transfer between languages.

In the end, both of these assessments are missed opportunities for teachers to learn about their bilingual students. Not bringing students' bilingualism to the forefront of analysis as a

primary lens in which understand student reading development, leaves the data collected, but only used minimally and, most likely, inaccurately.

## CHAPTER 5:

### Findings from the Surveys

In this chapter, the quantitative results from the surveys are presented. The survey developed for this research was designed to capture teacher self-reports about 1) how frequently they use assessments, 2) the type of knowledge that they gain from these, and 3) the instructional practices resulting from analyzing data from these assessments. As mentioned in the methodology section, the sample size was small (n= 63) and the response rate was low. All respondents were teachers of emergent bilingual students<sup>26</sup> who teach in grades 3-5. Respondents either filled out surveys at the Dual language Institute sponsored by the Department of Education in April 2011 or electronically through Survey Monkey. Both these factors – small sample size and low response rate – pose limitations on the generalizability of the findings. Nonetheless, the findings detailed here provide a complementary picture to the qualitative findings which describe how assessments impact teacher knowledge, and hence may impact reading instruction of emergent bilinguals.

In the qualitative chapter, I use data to support the argument that both types of assessments, summative and formative, fall short of their promise of providing teachers with knowledge that supports their teaching of emergent bilinguals. I also contend that the knowledge that is gleaned from these assessments is either not put into use or used minimally for the crafting of instructional practices for these students. In this chapter, I use the quantitative analyses to demonstrate that the patterns that were uncovered in the qualitative analysis reemerge

---

<sup>26</sup> Teachers who filled out the survey worked in three different types of programs: 64.15% worked in dual language bilingual programs, 22.64% in transitional bilingual programs and 13.21% in English as Second Language Programs.

upon analyzing the survey data, thus lending further evidence to the argument that both types of assessments studied do not contribute to the acquisition of “useable” knowledge by teachers of emergent bilinguals. The quantitative findings for the ELA corroborate the qualitative findings that most teachers of emergent bilinguals do not view these assessments as a source for the acquisition of knowledge about emergent bilinguals. In the case of running records, although teachers report learning from these assessments, the knowledge that they acquire from this assessment remains, for the most part, unused in the implementation of instructional practices best suited for emergent bilinguals. These quantitative findings add to the discussion of instructional practices that teachers implement after consulting data of emergent bilingual students that was started in the previous chapter. While the interviews documented teachers’ narratives about how data analysis is linked to instructional practices, the survey provided teachers with a standardized set of instructional practices. Therefore, the survey was able to collect information about a set of instructional practices.

The survey data was used to answer the following research questions. These questions overlap with the questions that were pursued in the qualitative analysis:

- 1) What kinds of knowledge do teachers of emergent bilinguals gain from the examination of reading assessment data of emergent bilinguals?
- 2) What changes, if any, in reading instructional practices for emergent bilingual students result from their teachers of emergent bilinguals’ *use* of large-scale standardized assessments and/or classroom-based assessments?
- 3) What changes, if any, in reading instructional practices for emergent bilingual students result from *the types of knowledge* gained from the use of large-scale standardized assessments and/or classroom-based assessments?

Before delving into the findings, it is important to highlight that the variables derived from the survey *quantify* teachers' self-reports about their experiences. Therefore, they represent respondents' approximations of the frequency with which they use an assessment, the extent of knowledge gained from assessment data, and the frequency with which they implement an instructional practice for emergent bilinguals as a result of assessment use and type of knowledge gained from assessment, rather than reflecting actual tallies of these. Furthermore, the variables in the study, with the exception of demographic variables, are *ordinal*. In other words, they represent a gradation in the frequency, or degrees, of how much an assessment used or instructional practice is used or the extent teacher knowledge is acquired from a particular assessment. These points are important to keep in mind, given that the results do not reflect degrees of how teachers actually learn or implement instructional practices, but rather an estimate of how much teachers perceive themselves as learning or changing their practices as a result of examining the assessment data of emergent bilingual students. In spite of the limitations of the research, these findings provide a feel for how teachers think of data from emergent bilinguals and the practices that they implement as a result of analyzing this data. The following section offers a description of how frequently teachers use the two assessments – the ELA and running records.

### **How frequently are ELA and Running Record Data Used?**

As described in Chapter 3, the survey collected data on how frequently teachers used four assessments (ELA, NYSESLAT, running records and conferencing). However, only data for the ELA and running records will be presented and contrasted in this chapter.

The question that was designed to collect this information specifically asked respondents how often they looked at assessments results to shape instruction *for emergent bilingual students*. Options ranged from “0 times” to “more than 8 times” per school year. Percentages of how often each of the assessments was used were calculated and are displayed in Tables 11 & 12. These percentages reflect the number of responses for a given frequency divided by the total number of responses collected for each question. For each of the tables, the total number of responses is listed as “n.” These tables provide readers with trends and the distribution of these variables.

Table 11 shows the percent of respondents who used the ELA to shape instruction for emergent bilingual students. The table reveals that the greatest reported use (33.9%) falls within the category “1-3 times per year.” In addition, close to 16% of the sample report not consulting data from this assessment. This finding was expected and reflects what was uncovered through teacher interviews that many teachers disregard the ELA data of their emergent bilingual students and therefore do not consult the results from this exam to shape instruction. Surprising was the finding that about 20% of the sample report consulting the ELA data of their emergent bilinguals “more than 8 times” a year. Furthermore, about 11% of teachers of emergent bilinguals reported using this assessment “7-8 times” a year. When these two frequencies are combined, it is revealed that about 1/3 of the teachers who answered this question consult the ELA over 7 times a year. This finding is unanticipated given the summative character of this assessment. In general, summative assessments are expected to be consulted at low frequencies since they provide teachers with a snapshot of student performance at the end of a given timeframe.

**Table 11. Percent Use of ELA Assessment (n=62)**

<b>Frequency</b>	<b>Percent of Respondents Who Use The Assessment at this Frequency</b>
0 Times	16.1
1-3 Times	33.9
4-6 Times	17.7
7-8 Times	11.3
More than 8 Times	21.0

The ELA is designed to provide an overview of students' grade-level abilities rather than information about student reading development along the year to aid in day to day instruction. As an indicator of student abilities at the end of the year, the ELA is most likely to be consulted as a "benchmark" in order to gain an understanding of where student literacy abilities may hover in relation to grade-level expectations. Benchmark assessment data, while consulted, is generally used at the end or the beginning of the year and not used with great frequency throughout the year. Used as such, it would be expected that the ELA would be consulted with less frequency than what was reported by teachers who answered this question. This finding fits into the qualitative findings which demonstrated that teachers use summative assessment data formatively.

Table 12 provides the percentage distribution for use of running records. The percent frequency for this assessment matches its formative purpose. Formative assessments are designed to provide teachers with ongoing information about student progress along the year. Therefore, it was expected that teachers would report consulting running record data of emergent bilinguals with great frequency. In fact, about 35% of teachers noted using running record data of emergent bilinguals "4-6 times" a year. When combined with other categories, "7-8 times"

and “more than 8 times” a year, it was found that 77.8% of teachers report using running record assessment data of emergent bilinguals at least “4-6 times” or more a year.

**Table 12. Percent Use of Running Records (n=63)**

<b>Frequency</b>	<b>Percent of Respondents Who Use The Assessment at this Frequency</b>
0 Times	7.9
1-3 Times	14.3
4-6 Times	34.9
7-8 Times	12.7
More than 8 Times	30.2

It is notable that close to 8% of teachers reported **not** using this form of assessment at all with emergent bilingual students. This finding is unusual since running record data is often collected by school administrators to monitor reading development of students school-wide. Furthermore, all of the teachers interviewed reported using running records. Perhaps, this disparity in results between the survey and the interviews is reflective of the difference in the sample composition between the interview and survey participants. While the interview sample was comprised entirely of dual language bilingual teachers, 11% of the teachers in the survey sample were ESL teachers. These teachers often do not head their own classroom, but rather work with small groups of students. It is possible that these teachers do not assess their students using running records.

The trends depicted in the tables above reflect the nature and purpose of the summative and formative assessments featured in this survey. It is of no surprise that the ELA, a summative assessment, was used with less frequency than the formative assessment, running records, based on the literature and previous findings from the interview data. This trend makes sense given that these types of summative assessments provide teachers of emergent bilinguals with data

about grade-level standards and attainment at the end of the school year, while formative assessments are used to inform teachers about emergent bilingual student development throughout the year. Furthermore, as stated beforehand, teachers interviewed lacked assurance that the assessment data of emergent bilinguals from the ELA would accurately reflect their abilities and therefore are more likely to disregard it. These initial trends confirm that teachers report using these two assessments differently depending on the nature of the assessment. In general, formative assessments, such as running records, are used more frequently than summative ones, such as the ELA.

Although these trends were expected, it is important to ask how these trends would hold in comparison to responses from a sample of teachers that did not teach emergent bilinguals. The patterns of assessment use seen here may be the same for teachers of students who are not emergent bilinguals. In that case, these trends may be reflective of how often teachers *in general* use assessments to shape instruction for their students. Although this question is beyond the scope of this data, it is important to keep this point in mind for the purpose of highlighting the limitations of this data as well as identifying avenues for further research. Furthermore, without the data on how frequently teachers use assessment data in general, regardless if their students are emergent bilinguals or not, comparisons cannot be made between the data on assessment use for emergent bilinguals and for “general education” students. Without these comparisons, possible reasons for the differences in use between assessments cannot be drawn.

Methodologically, it is important to note that the data for assessment use are not distributed normally. Therefore, non-parametric tests will be used when analyzing data

hereafter. The next section examines what types of knowledge teachers report acquiring from the data of emergent bilinguals from both the ELA and running records.

### **What Do Teachers Learn from the Assessment Data of Emergent Bilinguals?**

Teachers were asked the extent of knowledge they acquired from each of the assessments — the ELA and running records. The survey was designed to capture four types of knowledge that they may glean from these assessments: pedagogical content, student, assessment, and bilingual knowledge. The question in the survey was posed as, “When you look at your ELL students’ *English Language Arts Exam* how much do you learn about ...?” Another separate question asked about what teachers learned from examination of their emergent bilingual students’ running records results. Therefore, for a given type of knowledge, for example pedagogical content, there are two variables, one representing pedagogical content knowledge derived from the ELA and the other reflecting pedagogical content knowledge derived from running records.

Data from this question is displayed in two complementary ways. First, frequency distribution tables were constructed. These tables display three categories describing the extent of knowledge teachers report acquiring. The first category of “low” is the sum of teachers’ responses that they learned “not at all” and “a little.” The second category of “moderate” represents the choice offered to respondents, that they learned “a moderate amount.” The last category, “high,” represents the sum of teachers’ responses that they acquired “a lot” combined with “a great deal” of a given knowledge type. Frequencies are given in percentages. These percentages were calculated by taking the values for each of the original five frequencies

included in the survey and dividing them by the total number of responses collected for each question.

Secondly, cumulative distributive frequency graphs (hereafter referred to as CDFs) were created and placed alongside the frequency distribution tables. These graphs display cumulative percentiles of observations for a given variable (Glass & Hopkins, 1996: 24). The benefit of displaying data in this format is that it facilitates an understanding of the percent of respondents who answered at and below or above a certain value on the x-axis. In addition, the shape of the CDF line provides information about whether respondents selected higher or lower values for a given variable indicating whether teachers gained more or less levels of a given type of knowledge. For example in a CDF graph, a line that is curved upwards (with the hump on the top) indicates a greater percentage of respondents selected lower values for that variable. In this case, these lower values signify that teachers reported learning less of that type of knowledge.

In contrast, a line that is curved downwards (bowl-shaped) reveals that respondents provided higher values for that variable demonstrating that more teachers reported greater degrees of learning. The CDF format is suited to this analysis because it offers a clear comparison of the cumulative percentage of teachers' responses for variables mediated either through ELA or running record analysis. Furthermore, steep slopes in the CDF lines indicate where large changes in teachers' self-reporting are located. CDFs were constructed to compare how teachers reported acquiring four different types of knowledge either through the ELA or running records (RR).

To complete the analysis, the Wilcoxon signed rank test was used to determine whether statistically significant differences exist between a given knowledge type either acquired through

the ELA or running records (i.e. ELA pedagogical content knowledge compared to RR pedagogical content knowledge). The Wilcoxon signed rank test is non-parametric and therefore does not require that data are distributed normally, which is the case for the majority of the variables in this study (Glass & Hopkins, 1996: 12).

The hypotheses being tested with these data analyses are the following:

- 1) Teachers do report acquiring knowledge from the assessment used with emergent bilinguals.
- 2) Teachers report greater changes in knowledge from the use of formative data of emergent bilinguals.

### ***Pedagogical Content Knowledge***

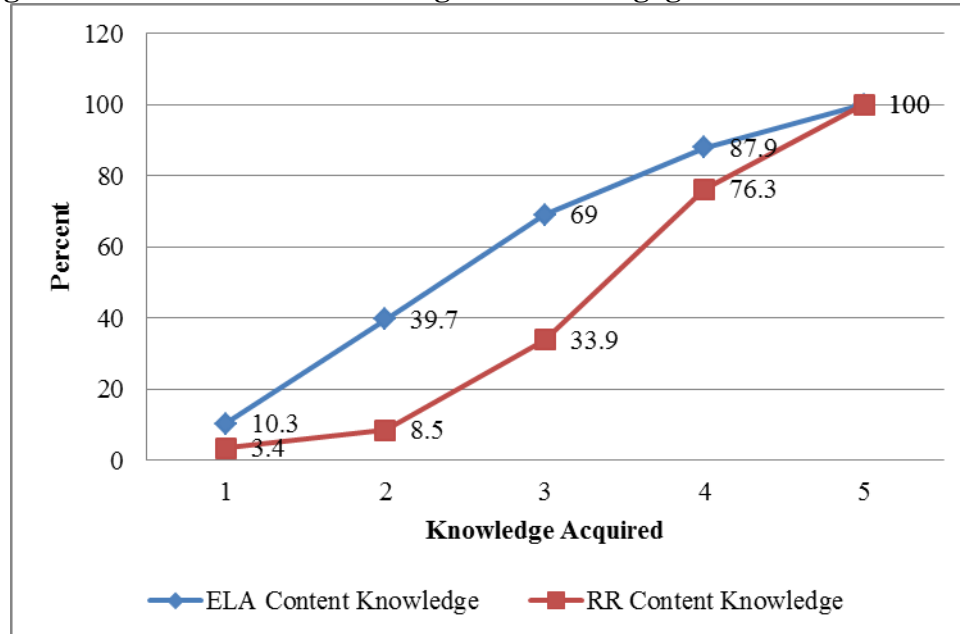
Pedagogical content knowledge was defined in this research as, the knowledge required for teachers to identify multiple ways to teach students in reading. Both Table 13 and Figure 1, which compare teachers' reported acquisition of pedagogical content knowledge either through the ELA or running records, indicate large differences between the two assessments in the reported acquisition of this type of knowledge. It was clear that the teachers did not report the ELA data of emergent bilinguals to be a source of pedagogical content knowledge. Nearly 40% of teachers reported acquiring low levels of pedagogical content knowledge after consulting ELA results. In contrast, about 66% of teachers stated acquiring high levels of pedagogical content knowledge from use of running records. Alternatively, only 31% of teachers reported high levels of pedagogical content knowledge from the ELA.

**Table 13. Frequency Distribution for Pedagogical Content Knowledge**

<b>Degree of Learning</b>	<b>ELA Pedagogical Content Knowledge (%) (n=58)</b>	<b>RR Pedagogical Content Knowledge (%) (n=59)</b>
Low	39.7	8.5
Moderate	29.3	25.4
High	31.0	66.1

Upon examining the CDF for pedagogical content knowledge, it is also important to point out the steep slope between “a moderate amount” (indicated by “3” on the x-axis) and “a lot” (indicated by “4” on the x-axis) for pedagogical content knowledge acquired through running records (see Figure 1). This slope points to a spike in responses at this level. Lastly, the line for pedagogical content knowledge acquired through running records is curved slightly downwards. This indicates that more teachers reported acquiring higher levels of knowledge than not. A Wilcoxon signed rank test showed that there was significant differences between teachers responses to both ( $Z = -3.8, p = .000$ ) signaling that teachers in this sample thought that they acquired more pedagogical content knowledge – or whether their reading strategies are effective for emergent bilinguals – from examining data from running records than from ELA data.

**Figure 1. CDF of ELA & Running Record Pedagogical Content Knowledge**



Note: 1 = Not at all; 2 = A Little; 3 = A Moderate Amount; 4 = A Lot; 5 = A Great Deal

### ***Student Knowledge***

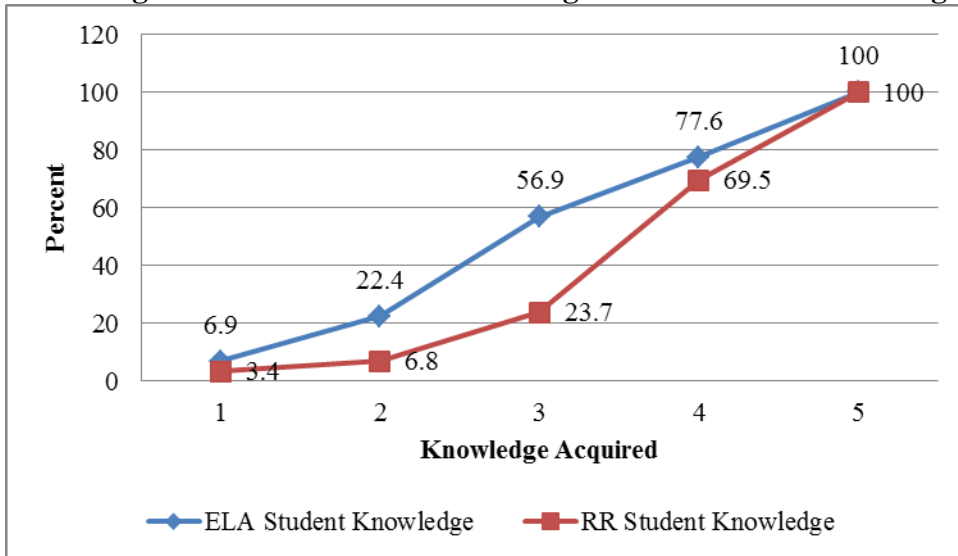
Student knowledge was defined in the survey as an understanding of “the skills and knowledge my ELL students possess.” Both Table 14 and Figure 2 indicate a similar trend for student knowledge acquired by teachers as a result of examining data of emergent bilinguals to that of pedagogical content knowledge. Teachers reported acquiring more student knowledge via running record data of emergent bilinguals than from ELA data of these students. As shown in Table 14, only 43.1% reported gaining high degrees of student knowledge of emergent bilinguals via ELA analysis in comparison to 76.3% of teachers reported gaining high levels of student knowledge through analysis of running records. In contrast, 22.4% of teachers reported that they acquired low levels of student knowledge via the ELA, compared to 6.8% of teachers who reported learning low levels of student knowledge from running records data of emergent bilinguals.

**Table 14. Frequency Distribution for Student Knowledge**

Degree of Learning	ELA Student Knowledge (%) (n=58)	RR Student Knowledge (%) (n=59)
Low	22.4	6.8
Moderate	34.5	16.9
High	43.1	76.3

Lastly, a Wilcoxon signed ranks test also demonstrated that there were significant differences between the two variables ( $Z = -3.172$ ,  $p = .002$ ). Teachers reported gaining more student knowledge about emergent bilinguals from running record analysis than from ELA analysis.

**Figure 2. CDF of ELA & Running Record Student Knowledge**



Note: 1 = Not at all; 2 = A Little; 3 = A Moderate Amount; 4 = A Lot; 5 = A Great Deal

### ***Assessment Knowledge***

Teachers reported acquiring more assessment knowledge from running records data of emergent bilinguals than from ELA data of these students (see Table 15 & Figure 3). About a third, 27.6%, of teachers believed that they gained a high degree of assessment knowledge about

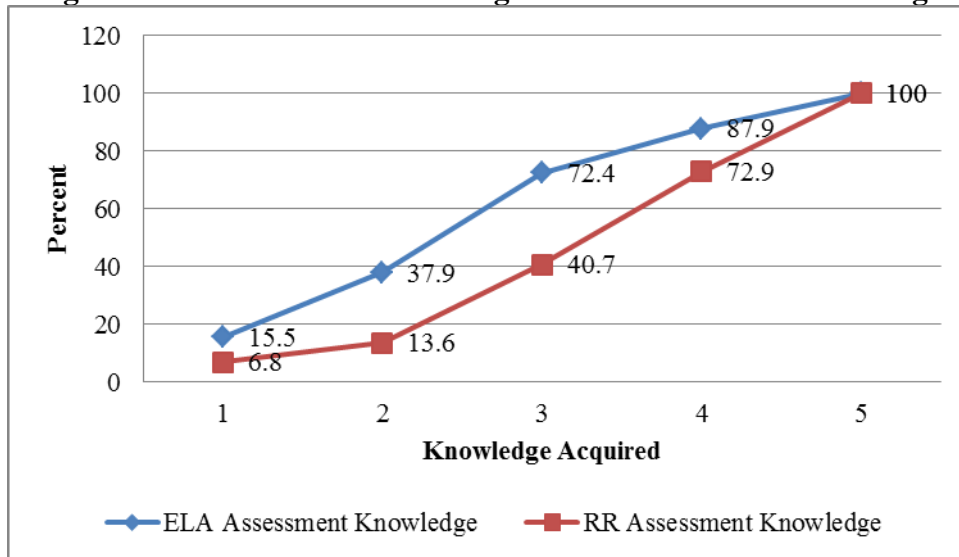
emergent bilinguals from analysis of the ELA. A greater percentage, 59.3%, of teachers reported gaining the same high levels of assessment knowledge through running records.

**Table 15. Frequency Distribution for Assessment Knowledge**

Degree of Learning	ELA Assessment Knowledge (%) (n=58)	RR Assessment Knowledge (%) (n=59)
Low	37.9	13.6
Moderate	34.5	27.1
High	27.6	59.3

The significance test used, Wilcoxon signed rank test, demonstrated significant differences in teacher reports between the two variables for assessment knowledge ( $Z = -3.691, p = .000$ ). Although teachers reported greater percentages of lower levels of acquisition of assessment knowledge than student knowledge for both ELA and running records, teachers reported acquiring more assessment knowledge from running records.

**Figure 3. CDF of ELA & Running Record Assessment Knowledge**



Note: 1 = Not at all; 2 = A Little; 3 = A Moderate Amount; 4 = A Lot; 5 = A Great Deal

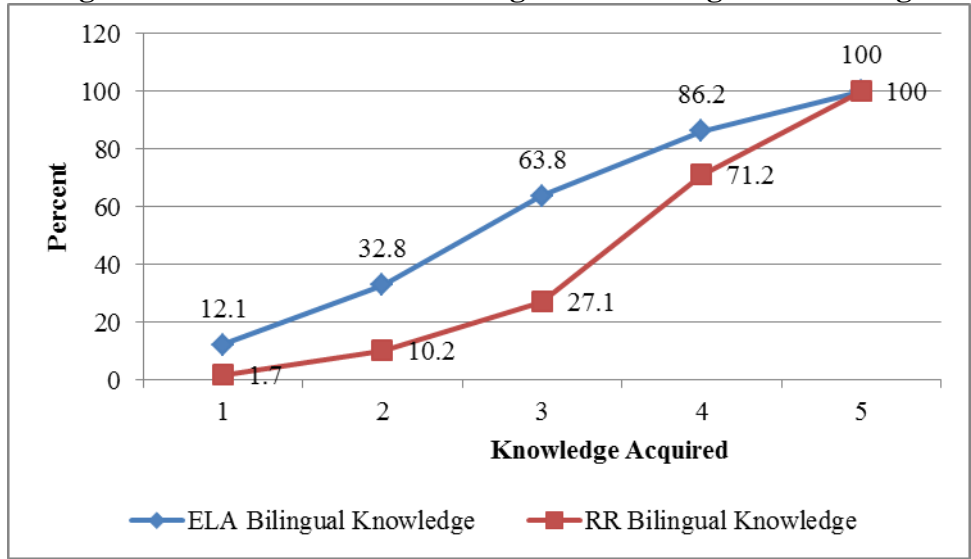
### ***Bilingual Knowledge***

Bilingual knowledge is defined as the understanding of how two or more languages develop and how to facilitate that development. With respect to the ELA analysis, about 36% of teachers reported that they acquired high degrees of bilingual knowledge after ELA analysis. In comparison 72.9% of teachers reported gleaning high levels of bilingual knowledge after analysis of running records. With regards to the ELA, 32.8% of teachers indicated that they acquired low levels of bilingual knowledge, whereas 10.2% of teachers reported acquiring the same low levels of knowledge after running record analysis. The Wilcoxon signed rank test also showed that these two variables differed significantly from each other ( $Z = -4.019$ ,  $p = .000$ ). Greater percentages of teachers reported acquiring higher levels of bilingual knowledge from running record data than from ELA data.

**Table 16. Frequency Distribution for Bilingual Knowledge**

<b>Degree of Learning</b>	<b>ELA Bilingual Knowledge (%) (n=58)</b>	<b>RR Bilingual Knowledge (%) (n=59)</b>
Low	32.8	10.2
Moderate	31.0	16.9
High	36.2	72.9

**Figure 4. CDF of ELA & Running Record Bilingual Knowledge**



Note: 1 = Not at all; 2 = A Little; 3 = A Moderate Amount; 4 = A Lot; 5 = A Great Deal

As was evident in the interviews, teachers of emergent bilinguals harbor a deep suspicion and reluctance to consider the ELA data of emergent bilinguals. These reservations are founded in the understanding that the ELA does not accurately assess many emergent bilingual students. The quantitative data show that teachers do not glean high levels of bilingual knowledge from the ELA. This may be one reason why teachers do not consult this data. In contrast, these findings show that teachers acquire high levels of bilingual knowledge from running records, although as was seen in the interviews this doesn't translate into adequate practices for emergent bilinguals.

### ***Type of Knowledge Gained***

These findings were also examined for differences in reported learning across the four types of knowledge. As shown in Table 17, a comparison of the percentile distributions of the extent to which teachers acquired knowledge from the ELA and running records revealed similar patterns. As we saw above, for all types of knowledge featured in this study a greater percentage

of teachers reported to acquire more knowledge about emergent bilinguals from running records than from ELA analysis. In fact, teachers consistently reported not learning at high levels after consulting the ELA data of emergent bilingual students. The Wilcoxon signed rank test, used to ascertain significant differences between the distributions for teacher knowledge acquired through ELA or running records, demonstrated that distributions for all types of teacher knowledge were significantly different.

Student knowledge was acquired at higher levels regardless of the type of data teachers consulted. In other words, a greater percentage of teachers reported gaining higher percentages of student knowledge *through* the ELA than for any of other type of knowledge acquired *through* the ELA. The same pattern holds for student knowledge acquired through running records – more teachers perceived to gain student knowledge *through* running records in comparison to other types of knowledge acquired *through* running records. It is expected that teachers report gaining student knowledge above all other types of knowledge through both of these assessments because this is the overarching “framework” in which teachers operate – assessments are primarily used to gain knowledge about students. Although I do not question this, without the complementary knowledge that content, assessment and bilingual knowledge provides, the student knowledge can remain inaccurate and generic.

**Table 17. Comparison of Frequency Distribution for Types of Teacher Knowledge**

Degree of Learning	ELA Pedagogical Content Knowledge (%)	RR Pedagogical Content Knowledge (%)	ELA Student Knowledge (%)	RR Student Knowledge (%)	ELA Assessment Knowledge (%)	RR Assessment Knowledge (%)	ELA Bilingual Knowledge (%)	RR Bilingual Knowledge (%)
Low	39.7	8.5	22.4	6.8	37.9	13.6	32.8	10.2
Moderate	29.3	25.4	34.5	16.9	34.5	27.1	31.0	16.9
High	31.0	66.1	43.1	76.3	27.6	59.3	36.2	72.9

From this chart it is also noteworthy that many teachers did not report acquiring assessment knowledge. One possible explanation is that assessment knowledge, the understanding of how assessments can lead to practices which support students, may be a concept that is new or nebulous to teachers. Generally, teachers are provided with assessments to use rather than seeing themselves as the authors or shapers of any given assessment. Therefore, while they administer an assessment or analyze the results, they may not engage in the type of thinking that enables them to develop sophisticated skills to change or adapt either the assessment or the process of analyzing to suit the characteristics of students. Furthermore, teachers may consider that assessment knowledge is acquired by other means, such as professional development, rather than through assessment itself.

Finally, both hypotheses are not rejected – 1) teachers do report acquiring knowledge from the assessment used with emergent bilinguals and 2) teachers report greater changes in knowledge from the use of formative data of emergent bilinguals – stated for this analysis.

The next section will explore if the extent to which teachers' frequency of use of either assessment is related to their acquisition of these types of knowledge. This question was not examined through the analysis of teacher interviews.

### **Is How Often an Assessment Used Related to Greater Levels of Acquisition of Teacher Knowledge?**

An additional step in this analysis was to examine how the patterns revealed through the prior analyses changed when examined alongside frequency of use of either assessment. This analysis looked into answering the following question: If teachers used an assessment more frequently, would this frequency of use be associated with higher levels of acquisition of teacher

knowledge? In order to answer this question, correlations were run between the variables for assessment use and teacher knowledge. The hypothesis being tested is as follows:

- 1) Reports of greater changes in teacher knowledge are positively associated with frequent use of assessment data.

Table 18, below, displays both the strength and the statistical significance between the ELA and running record use with the four types of knowledge. Moderate, positive correlations were found between running record use and all four types of knowledge acquired from running records. In contrast, there were no positive correlations between ELA use and teacher knowledge.

According to this analysis, teachers who reported using running records more frequently also tended to report acquiring greater levels of teacher knowledge. There were no statistically significant correlations between those teachers who used the ELA frequently and the acquisition of teacher knowledge.

**Table 18. Correlations between Assessment Use & Types of Teacher Knowledge**

	<b>Running Records Use (n=59)</b>	<b>ELA Use (n=57)</b>
<b>Pedagogical Content Knowledge</b>	.407***	.052
<b>Student Knowledge</b>	.424***	.140
<b>Assessment Knowledge</b>	.364***	.173
<b>Bilingual Knowledge</b>	.346***	.226

\*p < .05

\*\* p < .01

\*\*\* p < .001

Correlations do not imply causality nor do they indicate the direction of the relationship. Therefore, these findings cannot support the conclusion that teachers' frequency of assessment

use will impact the level acquisition of teacher knowledge. What can be derived from this analysis is that those teachers who use running records frequently are also more likely to report higher levels of learning from them or vice-versa (those teachers who report acquiring higher degrees of teacher knowledge are those who report using running records with greater frequency). A possible explanation for these results is that those teachers who already believe they learn a great deal from this particular assessment, use it more frequently than those who do not. In the case of this hypothesis – reports of greater changes in teacher knowledge are positively associated with frequency use of assessment data – it can be partially rejected. For the ELA it was found that there is no relationship between ELA use and the types of teacher knowledge studied. However, for running record use and types of teacher knowledge, there are statistically significant, moderate relationships and therefore, for this relationship the hypothesis is not rejected. The following analysis examines teachers’ self-reports about changes in instructional practice as a result of examining assessment data.

### **Do Teachers Report Changes in Instructional Practices as a Result of Examining Data?**

A central goal of this research was to examine if teacher knowledge mediates changes in instructional practices. Data for this analysis stemmed from the survey question which asked: “How often do you use the following instructional practices for ELLs *as a result of looking at* either the ELA or running records?”

As detailed previously, a menu of 17 instructional practices was provided to respondents through the survey (see Table 8). Furthermore, the individual practices featured in the survey can be grouped into the eight larger categories. The eight categories are:

1. grouping,
2. reading strategies,
3. work attack skills,
4. vocabulary,
5. background knowledge,
6. home language support,
7. oral language, and,
8. assessment.

Both frequency distribution tables and CDFs were constructed for each instructional practice used in the survey stemming either from an analysis of ELA or running record data of emergent bilinguals. A brief comparison of the two variables for a given instructional practice – one resulting from examining the ELA, one resulting from viewing running records – will be provided. The result of the Wilcoxon signed rank test is reported alongside these descriptions. Tables, charts, and the results of the statistical test are provided for each instructional practice. Furthermore, instructional practices are grouped into instructional categories in this section, in order to offer a summary of how teachers employ particular instructional strategies within a given category. The hypothesis being tested for all instructional practices is:

1. Teachers differ in the implementation of instructional practices as a result of either ELA or running record use.

***Instructional Category: Grouping***

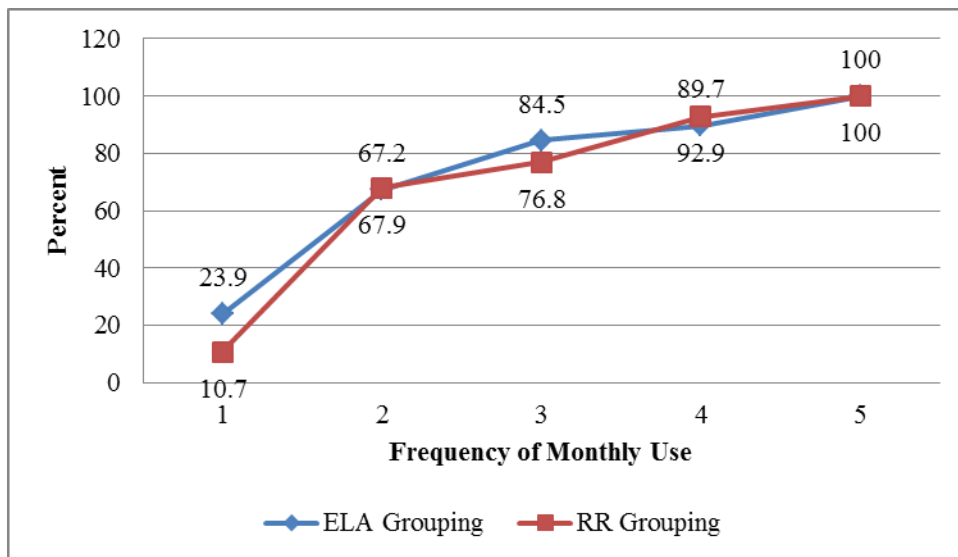
Grouping refers to when teachers, “regroup students into guided reading groups based on assessment data.” According Table 19 and Figure 5, the distributions for *grouping* as an instructional strategy upon consulting either ELA or running record data appear similar. In fact, about 67% of teachers reported low frequency of implementation for grouping for both assessments. According to the Wilcoxon signed rank test, these distributions are not

statistically significant from each other ( $Z = -1.062, p = .288$ ).

**Table 19. Frequency Distribution for Grouping as an Instructional Practice**

Frequency of Implementation	ELA Grouping (%) (n=58)	RR Grouping (%) <sup>27</sup> (n=56)
Low	67.2	67.9
Moderate	17.2	8.9
High	15.6	23.2

**Figure 5. CDF of Grouping as an Instructional Practice<sup>28</sup>**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

This finding stands in contrast to what was found through the interviews. A main finding from the interviews was that teachers reported using both ELA *and* running record data to form groups for reading instruction. The difference in results between the interviews and the survey may be due to the type of grouping teachers referred to. The survey specified teachers would use

<sup>27</sup> It is important to note that grouping refers to when teachers regroup students into guided reading groups. There are two variables that describe this one instructional practice, one that is the result of the ELA (ELA Grouping) and one that is the result of running records (RR Grouping). For the remainder of the instructional practices, for each singular instructional practice there are two variables – one as a result of the ELA and another as a result of running records.

<sup>28</sup> ELA Grouping refers to “grouping as a result of examining ELA data.” RR Grouping refers to “grouping as a result of examining running record data.” This notation will be used for all of the remaining instructional practices.

data to “regroup students into guided reading groups based on assessment data.” However, in interviews teachers referred to school-wide grouping rather than flexible grouping for guided reading that occurs periodically within classrooms. As proposed by the qualitative findings, as teachers go through the ritualized assessment practices in using assessment data, they are usually creating groups to receive support in creating school-wide groups, which is advocated by administrators, and not using this data to actually impact their day-to-day classroom grouping.

### ***Instructional Category: Reading Strategies***

The instructional practices of prediction, questioning, summarizing, and analysis of text all fall under the larger category of reading strategies. These instructional practices aid students in making meaning out of texts.

#### ***Prediction***

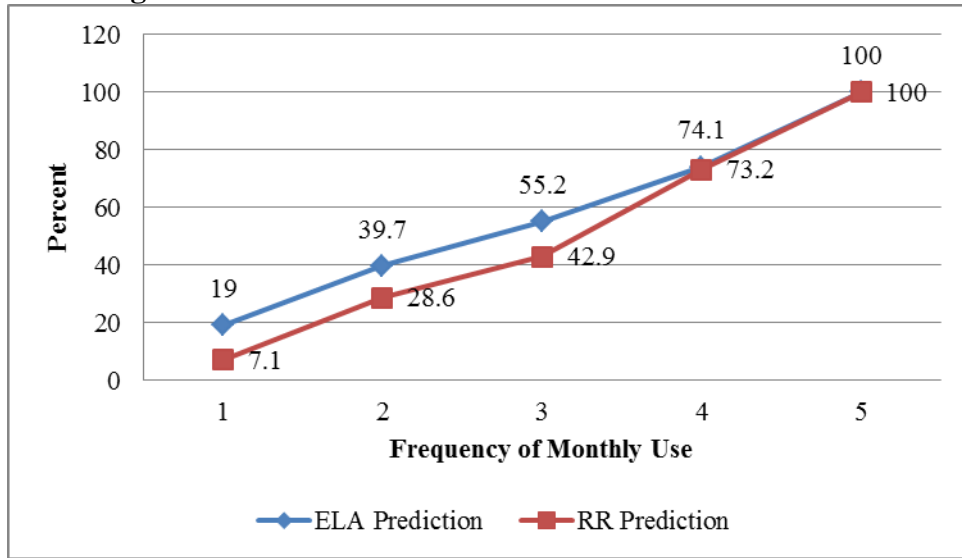
Prediction was defined in the survey as, “asking students to make predictions prior to reading.” This instructional strategy is important in helping students to prepare to read a given text successfully by activating their store of experiences that may relate to the book and ready their knowledge of vocabulary that is associated with the concepts in the text.

According to Table 20 and Figure 6, the distribution for *prediction* of an assessment strategy resulting from the examination of data from either ELA or running records are similar. Teacher reports of using this strategy are evenly distributed among the given frequencies. The Wilcoxon signed rank test confirmed that the distributions are not statistically significant from each other ( $Z = -1.718$ ,  $p = .086$ ).

**Table 20. Frequency Distribution for Prediction as an Instructional Strategy**

Frequency of Implementation	ELA Prediction (%) (n=58)	RR Prediction (%) (n=56)
Low	39.7	28.6
Moderate	15.5	14.3
High	44.8	57.1

**Figure 6. CDF of Prediction as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

### ***Questioning***

Questioning is when teachers ask students to generate their own questions about texts.

This reading strategy facilitates engagement of students with texts and offers students a tool with which to monitor their comprehension.

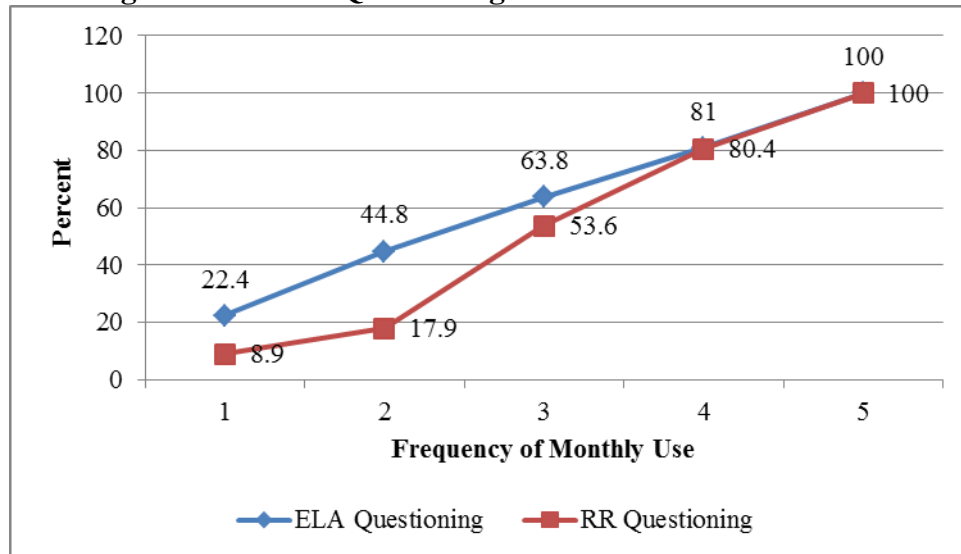
The percentile distributions for *questioning* as an instructional practice differs between ELA and running records (see Table 21 and Figure 7). This difference lies in the greater percentage of teachers who used questioning as an instructional practice as a result of examining running record data. Nearly 47% of teachers reported a high degree of the use of questioning as a result of using running records as compared to around 36% of teachers who used questioning at

this level as result of examining ELA data. The Wilcoxon signed rank test demonstrates that these distributions are statistically significant from each other ( $Z = -2.363$ ,  $p = .018$ ).

**Table 21. Frequency Distribution for Questioning as an Instructional Practice**

Frequency of Implementation	ELA Questioning (%) (n=58)	RR Questioning (%) (n=56)
Low	44.8	17.9
Moderate	19.0	35.7
High	36.2	46.4

**Figure 7. CDF of Questioning as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

### ***Summarizing***

Summarizing is an instructional practice which requires students to retell key points of a given text. Again, this strategy is meant to support student comprehension.

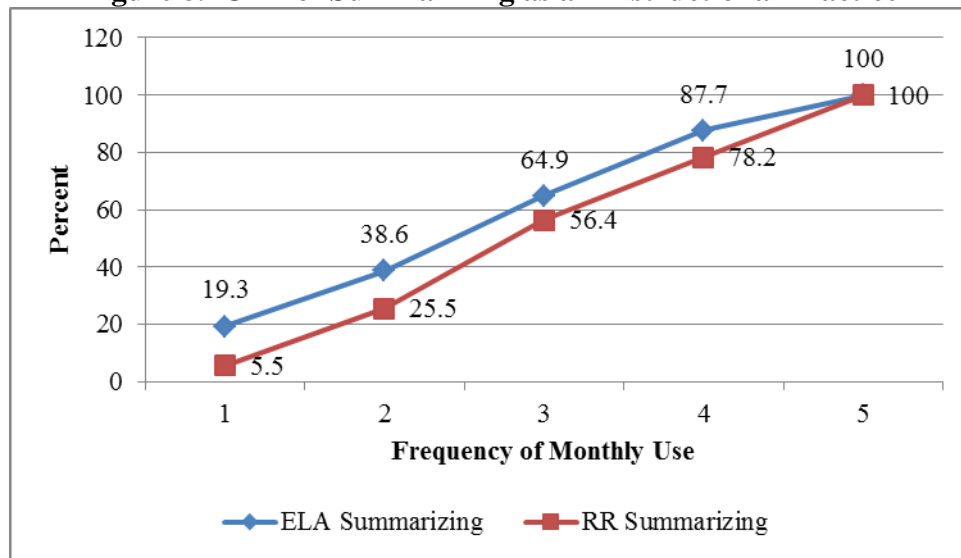
Although the CDF lines representing the percentile distribution for *summarizing* for ELA and running records appear similar, there are differences between teachers' responses for this instructional strategy. The main difference lies in the percentage of teachers who never use this instructional practice (see Figure 8). About 20% of teachers reported never using this instructional practice as a result of consulting the ELA exam, while the figure for teachers is

around 6% for teachers who examine running records data (see Figure 8). The Wilcoxon signed rank test confirms that these two distributions are significantly different from each other ( $Z = -2.363, p = .018$ ).

**Table 22. Frequency Distribution for Summarizing as an Instructional Practice**

Frequency of Implementation	ELA Summarizing (%) (n=57)	RR Summarizing (%) (n=55)
Low	38.6	25.5
Moderate	26.3	30.9
High	35.1	43.6

**Figure 8. CDF of Summarizing as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

### ***Analysis of Texts***

The analysis of texts is an instructional practice in which teachers ask student to offer their unique opinion or synthesize the ideas within a text. This is a higher order thinking skill that allows students to make meaning of texts.

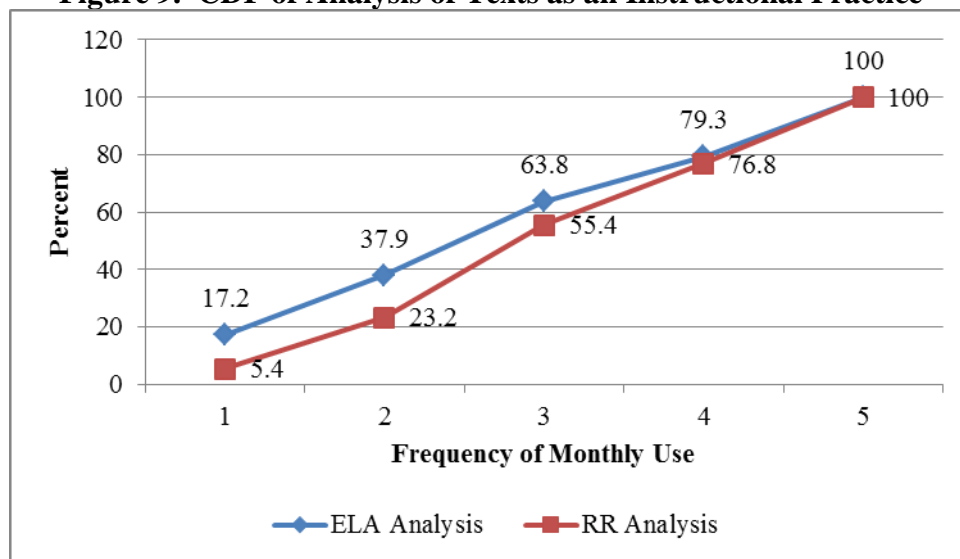
The data for the instructional practice, *analysis of text*, is similar to that for summarizing. Again the difference lies in the higher percentage of teachers who report never using this instructional strategy as a result of the ELA than after viewing running record data of emergent

bilinguals. According to the CDF, 44.6% of teachers reported high levels of usage of this strategy as result of looking at the running record test data and 36.2% of teachers reported the same percentage of usage after viewing ELA data. A Wilcoxon signed rank test signals that these distributions are significantly different from each other ( $Z = -2.018, p = .044$ ).

**Table 23. Frequency Distribution for Analysis of Texts as an Instructional Practice**

Frequency of Implementation	ELA Analysis of Texts (%) (n=58)	RR Analysis of Texts (%) (n=56)
Low	37.9	23.2
Moderate	25.9	32.2
High	36.2	44.6

**Figure 9. CDF of Analysis of Texts as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

***Summary: Reading Strategies Category***

In sum, Wilcoxon signed rank tests demonstrated that the distributions for questioning, summarizing and analyzing texts were significantly different from each other when using the ELA and the running record. A greater percentage of teachers reported implementing all types of reading strategies, including prediction, for emergent bilinguals after consulting running record data than ELA data. However, these strategies were employed with high frequency (“12-

16 times a month” to “everyday”) less than 50% of the time after consulting either the ELA *or* running records. Therefore, although there are differences between implementation of these strategies after examining running record data over ELA data, these type instructional practices are not practiced with notable frequency after examination of either type of data for emergent bilingual students.

***Instructional Category: Word Attack Skills***

The teaching of sight words and decoding skills can be grouped under the larger category of instructional practices considered “word attack skills.” Word attack skills are necessary for students to access words in texts and read fluently.

***Sight Words***

Sight words are frequently used words that students should recognize and be able to read automatically. Having sight word recognition is important in reading texts fluently.

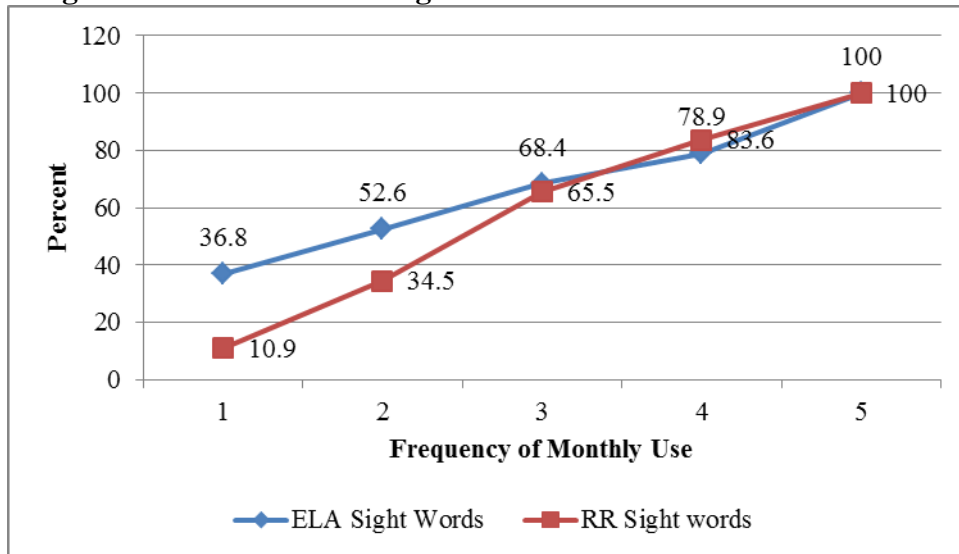
A Wilcoxon signed rank test shows that teachers’ reports for use of the instructional practice of teaching *sight words* after examining data from the ELA or the running records do not differ significantly from each other ( $Z = -1.937$ ,  $p = .053$ ). About the same percentage of teachers report teaching sight words with a high frequency of implementation (31.6% after consulting the ELA and 34.6% after viewing running records results, see Table 24). Despite this, there are some differences. A higher percentage of teachers report never implementing this practice after consulting the ELA, 36.8%, compared to 10.9% for running records (see Figure 10). What this finding points to is that teachers use this instructional practice, the teaching of sight words to emergent bilingual students, at similar rates regardless of the assessment used. It can be conjectured therefore that this instructional practice is independent of assessments, and is

implemented by teachers because it is within their arsenal of instructional practices for use with their emergent bilingual students.

**Table 24. Frequency Distribution for Sight Words as an Instructional Practice**

Frequency of Implementation	ELA Sight Words (%) (n=57)	RR Sight Words (%) (n=55)
Low	52.6	34.5
Moderate	15.8	30.9
High	31.6	34.6

**Figure 10. CDF of Teach Sight Words as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

### ***Decoding***

Decoding refers to the ability of students to use phonemic awareness in order to say and hence understand the written word. Decoding is a critical skill for students in order to figure out unknown words that are written but are within their store of oral language vocabulary.

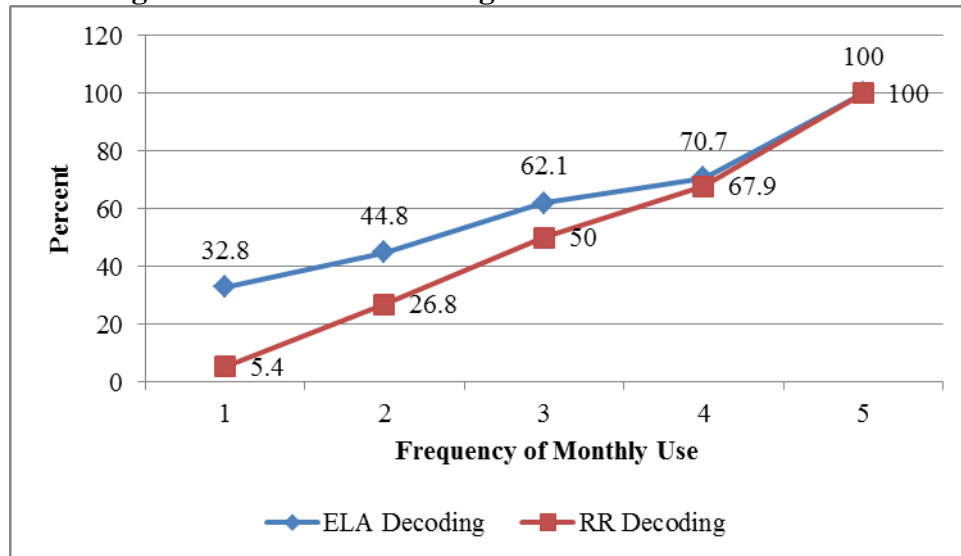
The teaching of *decoding* as an instructional practice differs significantly after either examining ELA and running record data, according to the Wilcoxon signed rank test ( $Z = -2.392$ ,  $p = .017$ ). Fifty percent of teachers report using decoding as an instructional practice after

consulting running record data with high frequency (see Table 25). Almost 6 times more teachers stated that they never use this strategy after viewing ELA results (see Figure 11).

**Table 25. Frequency Distribution for Decoding as an Instructional Practice**

Frequency of Implementation	ELA Decoding (%) (n=58)	RR Decoding (%) (n=56)
Low	44.8	26.8
Moderate	17.3	23.2
High	37.9	50.0

**Figure 11. CDF of Decoding as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

***Summary: Word Attack Instructional Category***

To sum up this section, there were only two types of word attack instructional practices featured in the survey. For the teaching of sight words as an instructional practice, there was no difference in response after viewing either ELA or running record data. In fact, less than 40% of teachers reported teaching sight words with high frequency after viewing data of emergent bilinguals from either assessment. This may be attributable to the use of sight words in earlier grades, K-2, rather than in grades 3 and higher. Although there were differences in the

implementation of decoding after ELA and running record data analysis, the differences do not appear to be large.

***Instructional Category: Vocabulary***

Three instruction practices – use of texts with limited vocabulary, teaching students required vocabulary, and teaching students to figure out the meaning of unknown words – were featured in the survey. These three instructional practices can be categorized as “vocabulary” instructional practices.

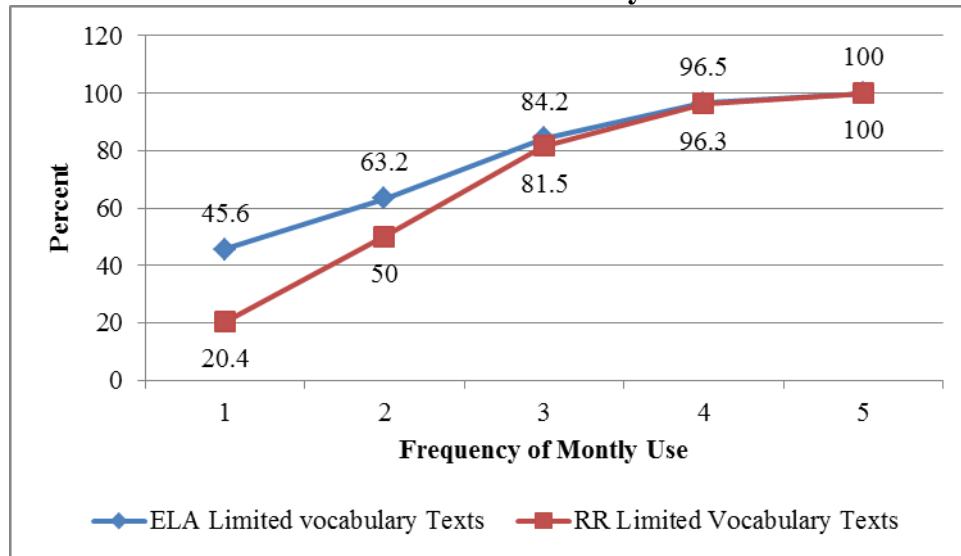
***Texts with Limited Vocabulary***

Using texts with limited vocabulary is an instructional practice that teachers may implement to control the level of texts that students have access to. Table 26 demonstrates that higher rates of teachers report never using the instructional practice of *using texts with limited vocabulary* with emergent bilinguals after viewing ELA results (see Figure 12). While there are clear differences between use of this practice after consultation of ELA and running record use, the percentages of teachers who use this instructional practice infrequently are much higher than other practices seen thus far. A Wilcoxon signed rank test shows that the distributions differ statistically ( $Z = -2.129, p = .033$ ).

**Table 26. Frequency Distribution for Texts with Limited Vocabulary as an Instructional Practice**

<b>Frequency of Implementation</b>	<b>ELA Texts with Limited Vocabulary (%) (n=57)</b>	<b>RR Texts with Limited Vocabulary (%) (n=54)</b>
Low	63.2	50.0
Moderate	21.1	31.5
High	15.7	18.5

**Figure 12. CDF of Texts with Limited Vocabulary as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

### ***Teaching Students Required Vocabulary***

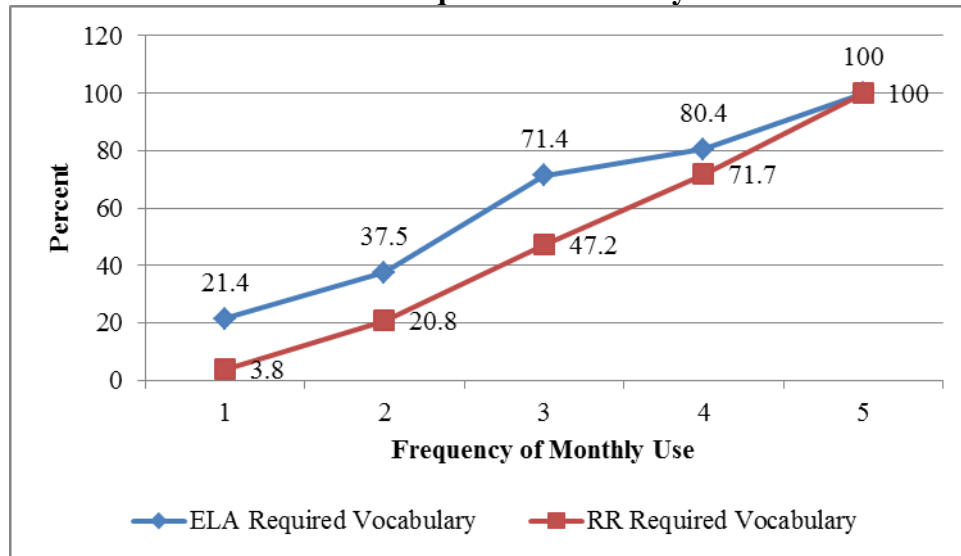
Teaching students required vocabulary is an instructional practice that teachers may implement in order to provide students with a “headstart” in tackling tricky vocabulary within a text. Usually teachers will introduce words that they anticipate students not knowing before students read the given text.

The percentile distributions for *teaching required vocabulary* clearly differ (see Table 27 and Figure 13). A higher percentage of teachers report a high frequency of implementation after viewing running record data of emergent bilinguals. The Wilcoxon signed rank test shows that indeed these distributions differ significantly from each other ( $Z = -2.802, p = .005$ ).

**Table 27. Frequency Distribution for Teaching Students Required Vocabulary as an Instructional Practice**

Frequency of Implementation	ELA Teach Students Required Vocabulary (%) (n=56)	RR Teach Students Required Vocabulary (%) (n=53)
Low	37.5	20.8
Moderate	33.9	26.4
High	28.6	52.8

**Figure 13. CDF of Teach Students Required Vocabulary as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

### ***Teaching Meaning of Unknown Words***

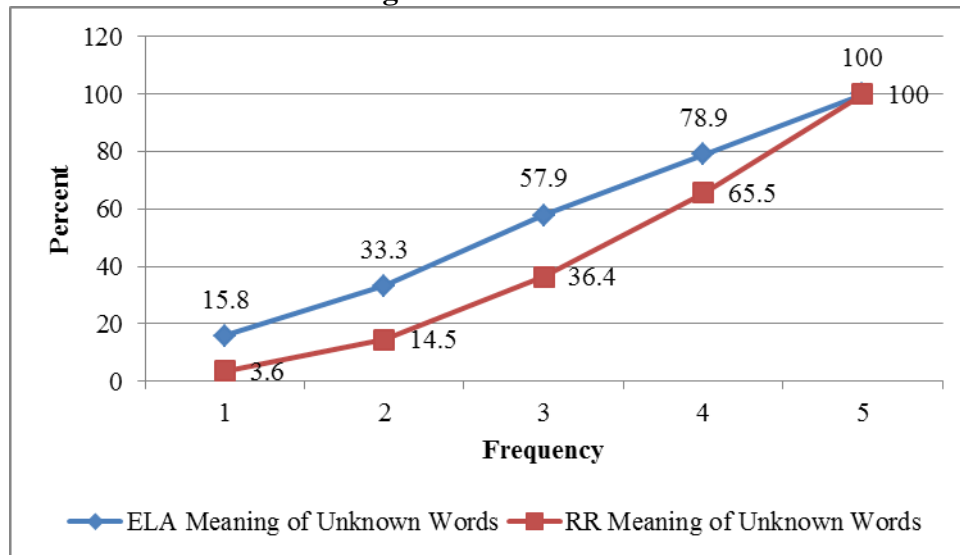
This instructional practice was defined in the survey as, “teach students how to figure out the meaning of unknown words (ex. synonyms, context clues, and cognates).” Teaching students how to use their word knowledge in order to figure out other known words is a vital skill that contributes to widening their vocabulary. In addition, it provides students with the opportunity to figure out the meaning of words while they read, rather than seeking help elsewhere.

Table 28 and Figure 14 show that a higher percentage of teachers report using this instructional practice with high frequency after viewing running record results (63.7%) than after viewing ELA results (42.1%). The test of significance demonstrates that this difference is statistically significant ( $Z = -2.637, p = .008$ ).

**Table 28. Frequency Distribution for Teach Meaning of Unknown Words as an Instructional Practice**

Frequency of Implementation	ELA Teach Meaning of Unknown Words (%)	RR Teach Meaning of Unknown Words (%)
Low	33.3	14.5
Moderate	24.6	21.8
High	42.1	63.7

**Figure 14. CDF of Teach Meaning of Unknown Words as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

**Summary: Vocabulary Instructional Category**

The comparison of percentage frequency of implementation of vocabulary practices yielded some interesting results. Teachers overwhelmingly did not use the practice of using *texts with limited vocabulary* with high frequency (“12-16 times a month” to “every day”) after viewing either ELA or running record data. In contrast, teachers reported using both *teaching*

*required vocabulary and teaching meaning of unknown words* with high frequency (“12-16 times a month” to “every day”) after viewing running record data. While teachers opt not to use texts with limited vocabulary with their emergent bilingual students after consulting data in general, teachers do report teaching both required vocabulary and the meaning of unknown words at high rates after consulting running record data. This signals that teachers identify the need to teach emergent bilingual students both actual vocabulary and vocabulary skills in order to facilitate their reading. This result matches the qualitative finding that teachers mentioned running records as a source to identify vocabulary words and skills that students would need. However, what was found through the interviews is that teachers loosely based their reasoning about vocabulary on running record data. Rather, their implementation of this instructional practice was founded on existing knowledge that emergent bilingual students need vocabulary instruction in order to be successful readers. What this quantitative result suggests is that teachers most likely use both — existing knowledge and information gleaned from the running record — to determine the vocabulary and related skills their emergent bilingual students need.

***Instructional Category: Background Knowledge***

Two instructional practices featured in the survey can be categorized under the larger group of instructional practices known as “background knowledge.” These are “activating prior knowledge” and “providing students with new background knowledge.” The data for these instructional practices will be described together because of their similarities (see Tables 29 & 30 and Figures 15 & 16).

### ***Activating Prior Knowledge & Providing Students with Background Knowledge***

Activating prior knowledge is an instructional practice in which teachers ask questions, use anticipation guides or questionnaires in order to “set the stage” for students to read. It allows students to make deeper connections to the text and thus generally improves comprehension. Providing students with background knowledge is complementary to activating prior knowledge. If students lack the experience or knowledge of a given topic covered in a text, teachers can do an activity, bring in materials or read introductory material in order to provide students with new background knowledge on a given topic.

For both of these instructional practices, more than 60% of teachers reported using them with high frequency. Wilcoxon signed rank test demonstrated that the distributions for activating prior knowledge were statistically significant from each other ( $Z = -2.243$ ,  $p = .025$ ). However, the distribution for providing background knowledge was not significant ( $Z = -1.885$ ,  $p = .059$ ). Although similar in distributions, this difference in significance is most likely due to the small sample size.

**Table 29. Frequency Distribution for Activate Prior Knowledge as an Instructional Practice**

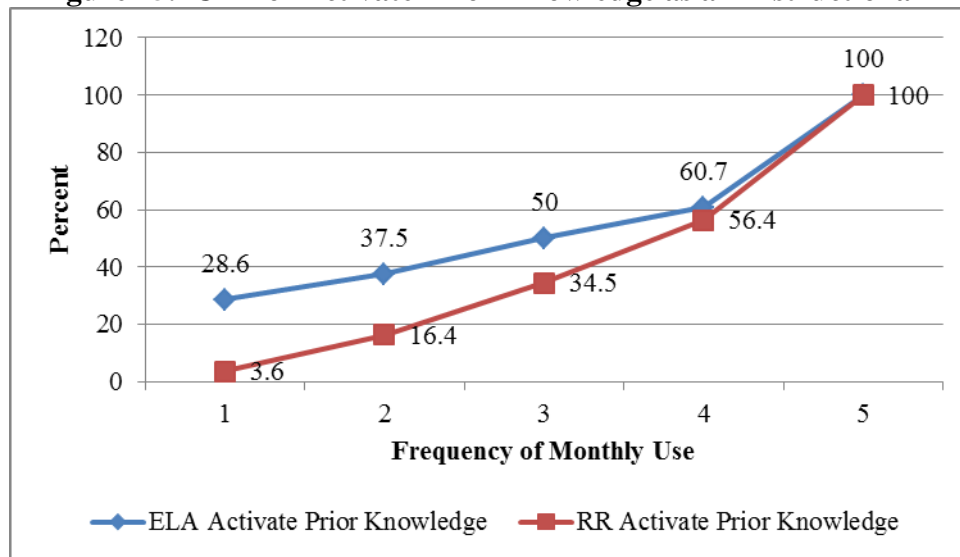
<b>Frequency of Implementation</b>	<b>ELA Activate Prior Knowledge (%) (n=56)</b>	<b>RR Activate Prior Knowledge (%) (n=55)</b>
Low	37.5	16.4
Moderate	12.5	18.2
High	50.0	65.4

**Table 30. Frequency Distribution for Provide Students with Background Knowledge as an Instructional Practice**

Frequency of Implementation	ELA Provide Students with Background Knowledge (%) (n=57)	RR Provide Students with Background Knowledge (%) (n=55)
Low	36.8	16.4
Moderate	12.3	21.8
High	50.9	61.8

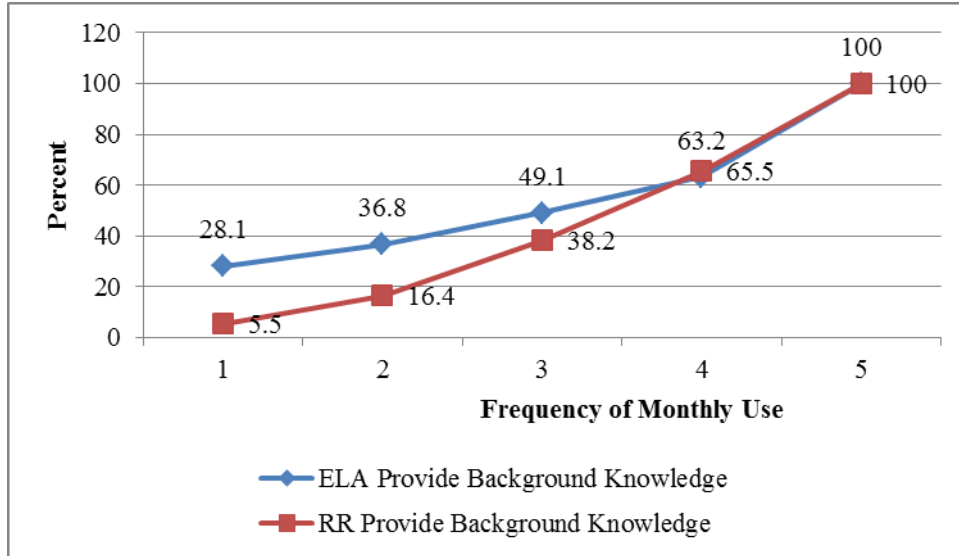
For both these instructional practices, teachers are more likely to implement instructional practices which supported emergent bilinguals’ background knowledge after consulting running record data rather than ELA data. Furthermore, out of all of the instructional practices seen thus far, these two are used by a far greater percentage of teachers as compared to other practices. In fact, more than 60% of teachers use both these “background knowledge” strategies with high frequency after consulting running record data of emergent bilingual students.

**Figure 15. CDF of Activate Prior Knowledge as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

**Figure 16. CDF of Provide Students with Background Knowledge as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

***Summary: Background Knowledge Category***

Teaching students through background knowledge – specifically by activating prior knowledge and providing new background knowledge to students – was frequently implemented by teachers after viewing running record results as well as after consulting the ELA, although to a lesser extent. This finding can be interpreted as teachers viewing this instructional strategy as a critical one to emergent bilingual students’ reading success. Although there are differences in the percentages to which this instructional strategy is implemented between the ELA and running records, teachers generally implement these strategies at high levels regardless of the assessment. This may mean that teachers partially use assessment results to determine whether to use this practice and implement it in combination with their existing knowledge.

***Instructional Category: Use of Home Language***

One instructional practice which highlights the use of home language, “provide students with texts in native language,” was featured in the survey. The use of texts in the home

language is an instructional practice which acknowledges that bilingual students may need support in their home language in order to advance their reading skills in their second one.

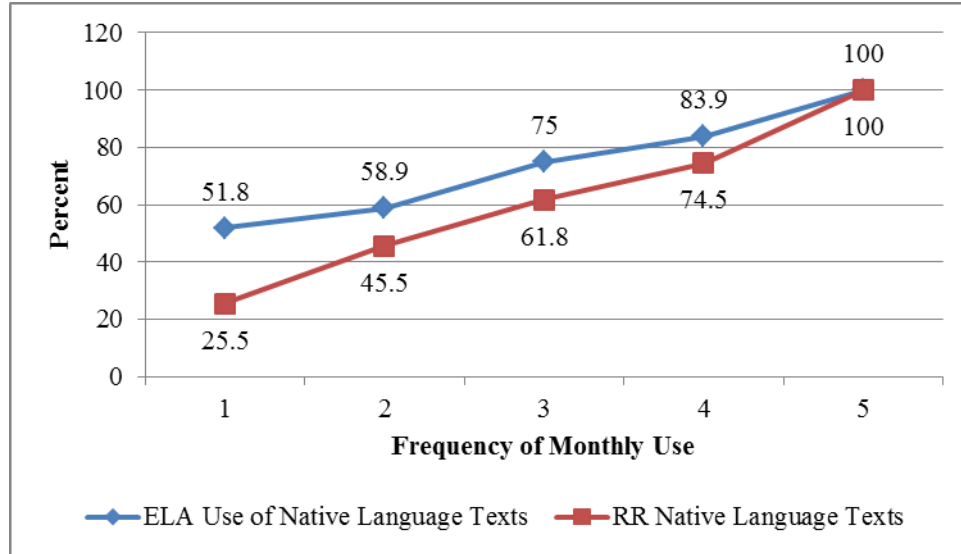
Some teachers do report using this strategy after examining results from either assessment (see Table 31). Twenty-five percent of teachers report using native language texts with high frequency after viewing ELA results, while 38.2% of teachers report implementing this instructional practice with the high frequency after viewing running record results. One remarkable thing evident from the data is that a great number of teachers never use this practice after either consulting ELA or running record results. According to the CDF, 51.8% of teachers never use this for emergent bilinguals after viewing ELA results, while 25.5 percent of teachers never use it upon examining running record results (see Figure 17). A Wilcoxon signed rank test demonstrated that these distributions differ significantly from each other ( $Z = -2.656, p = .008$ ).

Although there are differences in the implementation of this strategy after consulting ELA or running record data, the percentage distribution shows that many teachers do not use home language texts as an instructional practice at all.

**Table 31. Frequency Distribution for Provide Students with Texts in Native Language as an Instructional Practice**

<b>Frequency of Implementation</b>	<b>ELA Provide Students with Texts in Native Language (%) (n=56)</b>	<b>RR Provide Students with Texts in Native Language (%) (n=55)</b>
Low	58.9	45.5
Moderate	16.1	16.3
High	25.0	38.2

**Figure 17. CDF of Provide Students with Texts in Native Language as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

***Summary: Use of Home Language***

This result indicates that teachers largely do not access emergent bilingual students’ home language in order to further these students’ reading development. Despite this being a promising practice that in fact could aid in students’ reading progress, perhaps, this finding is due to teachers’ preconceived notions that they must stick to reading instruction in one language as was hinted in the qualitative findings. According to interviews, dual language bilingual teachers generally believe that they must separate English and Spanish reading instruction, and their use of data generally is skewed towards the importance of English reading, rather than Spanish reading. These qualitative results confirm this.

***Instructional Category: Oral Language***

Two instructional practices categorized under “oral language” were included in the survey. These are: “asking students to individually speak about reading” and “engaging students in partner talk.”

### ***Asking Students to Speak about Texts & Engaging Students in Partner Talk***

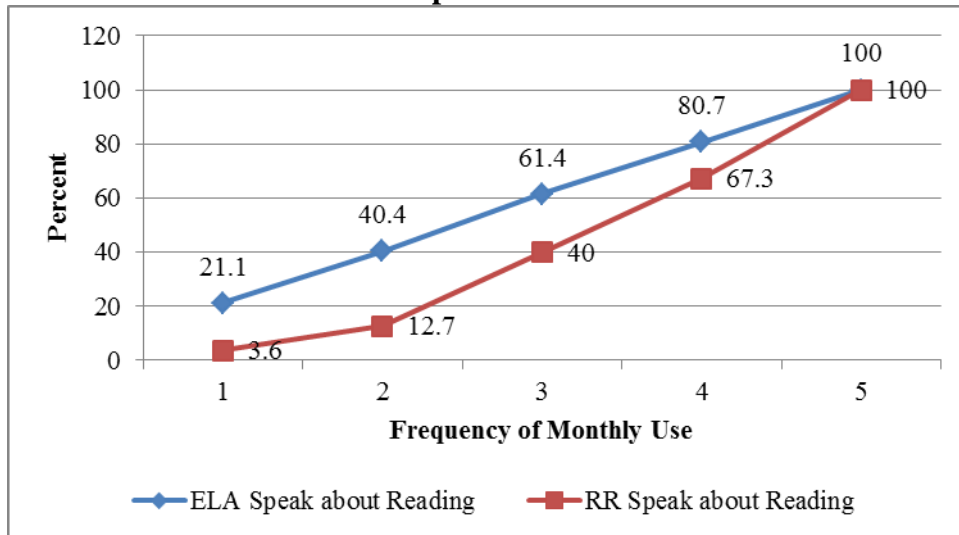
Oral language is important in developing literacy skills. The instructional practice of asking students to speak about texts allows students to answer questions and retell their understandings orally, rather than in writing. The strategy engages students in partner talk and allows them to talk to a peer about their questions, noticings and ideas about a text.

The two strategies, *asking students to speak about texts* and *engage students in partner talk*, demonstrated significant differences between use after consulting either ELA or running record data. For both of these instructional practices, teachers exhibited high frequencies of implementation (over 60%) after viewing running records as opposed to after consulting ELA data (lower than 40%). The Wilcoxon signed rank test showed that teachers responses about asking students to speak about texts as an instructional practice differed after viewing the ELA or running record results ( $Z = -3.231, p = .001$ ). The same was true for using partner talk as an instructional practice ( $Z = -2.025, p = .043$ ).

**Table 32. Frequency Distribution for Ask Students to Speak about Texts as an Instructional Practice**

<b>Frequency of Implementation</b>	<b>ELA Ask Students to Speak about Texts (%) (n=57)</b>	<b>RR Ask Students to Speak about Texts (%) (n=55)</b>
Low	40.4	12.7
Moderate	21.0	27.3
High	38.6	60.0

**Figure 18. CDF of Ask Students to Speak about Texts as an Instructional Practice**



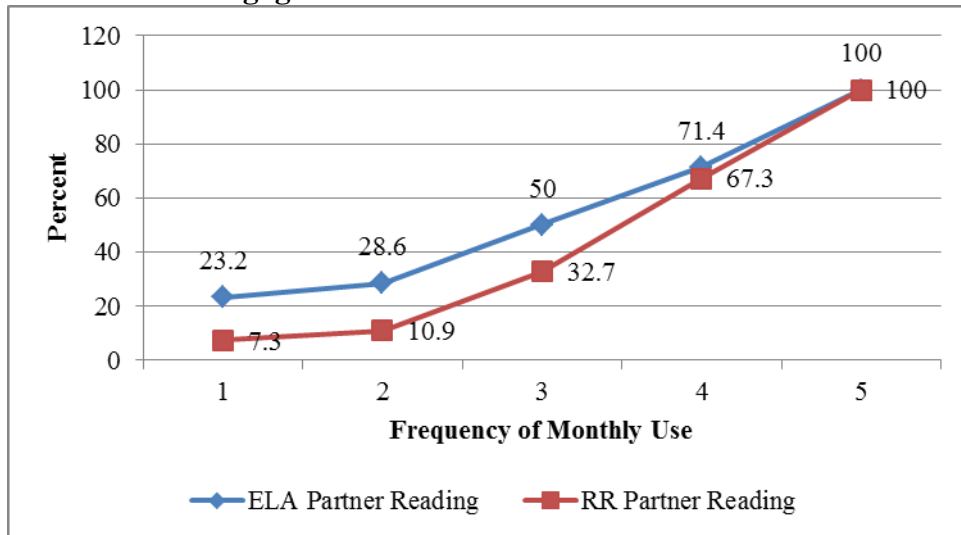
Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

The results for the instructional practice, *engaging students in partner talk*, differed from asking students to talk about what they read. In fact, a lesser percentage of teachers reported using this strategy never or infrequently (“1-3 times a month”) after either viewing ELA or running record results – 28.6% for the ELA and 10.9% for running records.

**Table 33. Frequency Distribution for Engage Students in Partner Talk as an Instructional Practice**

Frequency of Implementation	ELA Engage Students in Partner Talk (%) (n=56)	RR Engage Students in Partner Talk (%) (n=54)
Low	28.6	10.9
Moderate	21.4	21.8
High	50.0	67.3

**Figure 19. CDF of Engage Students in Partner Talk as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

***Summary: Oral Language***

In sum, the use of oral language instructional practices for emergent bilinguals differed after viewing either ELA or running record data. In fact, over 60% of teachers reported using these practices with great frequency after consulting running record data. These findings show that teachers overwhelmingly turn to providing students with opportunities for oral language practice. Furthermore, the levels at which teachers responded to implement these strategies after viewing running record data is far higher than after consulting ELA data. These findings suggest that running records results are indeed one factor that teachers consider as they formulate their instructional practices.

***Instructional Category: Assessment***

Two instructional practices categories under assessment were included in the survey. These were “providing students with opportunities to practice assessment” and “observe and

document” students’ reading (see Tables 32 & 33). These two practices overlap both instruction and assessment.

***Providing Students with Opportunities to Practice Assessment***

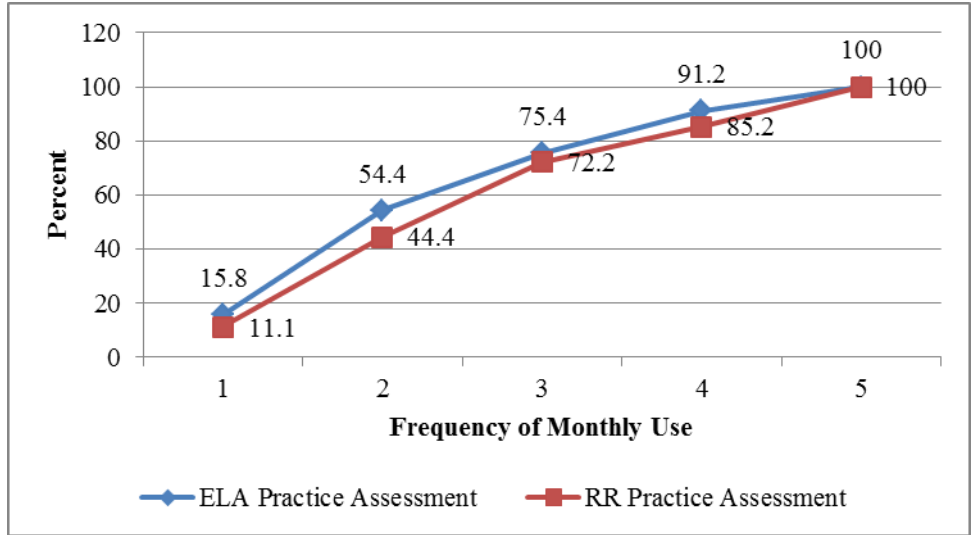
The first practice, *providing students with opportunities to practice assessment*, was featured in the survey because assessment results often led to more practice of that very assessment. This was my experience in schools.

Table 34 demonstrates that teachers overwhelming do not report providing further opportunities to practice an assessment after either viewing ELA or running record data. A Wilcoxon signed rank test shows no statistically significant differences between the two distributions ( $Z = -.0859$ ,  $p = .391$ ).

**Table 34. Frequency Distribution for Provide Students with Opportunities to Practice Assessment as an Instructional Practice**

<b>Frequency of Implementation</b>	<b>ELA Provide Students with Opportunities to Practice Assessment (%) (n=57)</b>	<b>RR Provide Students with Opportunities to Practice Assessment (%) (n=55)</b>
Low	54.4	44.4
Moderate	21.0	27.8
High	24.6	27.8

**Figure 20. CDF of Provide Students with Opportunities to Practice Assessment as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

***Observing and Documenting Students’ Reading as an Instructional Practice***

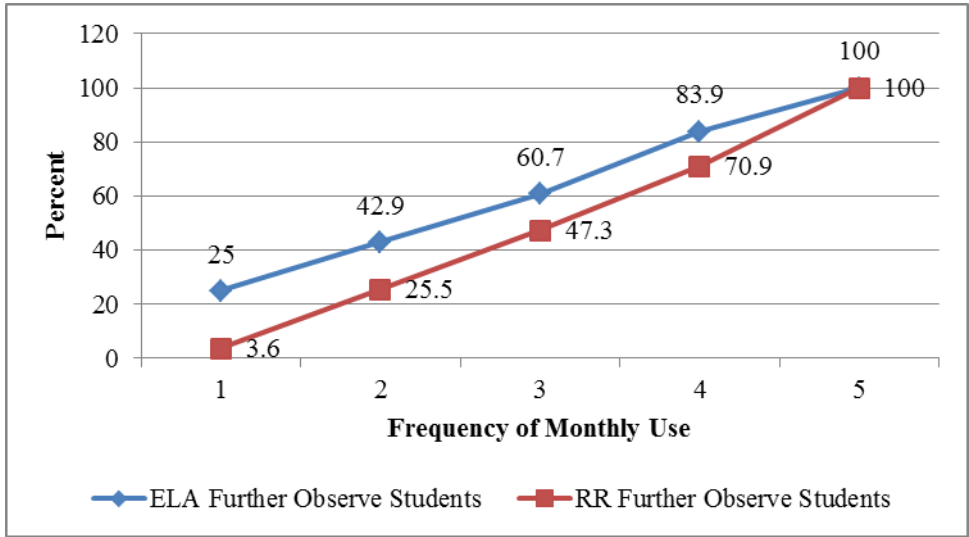
The instructional practice of observing and documenting student reading includes doing the work of “kid watching.” This combined instructional practice and assessment is the epitome of formative assessment, in that it allows teachers to access how students learn and approach reading by documenting students while they work. This instructional practice was included because it serves as a counterpoint the previous instructional practice of providing more assessment as an instructional practice.

This practice, *further observation and documentation of students*, showed a statistically significant difference between percentages of teachers who implemented this instructional practice after consulting ELA or running record data ( $Z = -2.453, p = .014$ ). A greater percentage of teachers, 52.7%, used this practice with high frequency after viewing running record results of emergent bilingual students (see Table 35). In contrast, 39.3% of teachers used this strategy with high frequency after consulting ELA data of these students.

**Table 35. Frequency Distribution for Observe Students as an Instructional Practice**

Frequency of Implementation	ELA Observe Students (%) (n=56)	RR Observe Students (%) (n=55)
Low	42.9	25.5
Moderate	17.8	21.8
High	39.3	52.7

**Table 21. CDF of Observe Students as an Instructional Practice**



Note: 1 = Never; 2 = 1-3 Times a Month; 3 = 4-8 Times a Month; 4 = 12-16 Times a Month; 5 = Every Day

***Summary of Assessment Instructional Practices***

The findings for this category of instructional practices suggest that teachers do not decide to provide students with opportunities to practice either assessment after viewing results from either. Although teachers suggested in the interviews that they do test-prep in order to expose and ready their emergent bilingual students for the ELA, these findings suggest that they do so not as a result of viewing data from this assessment, but rather as part of the larger framework of assessment behaviors in which they operate. However, teachers did report researching their emergent bilingual students further through additional observation after consulting running record data. This result may signal that teachers in fact partially base their decision to do additional formative assessment based on running record data of their students.

## Summary of Analysis of Instructional Practice

Out of the 17 instructional practices examined through this survey, an analysis of the percentile distributions and examination of Wilcoxon signed rank test signaled that teachers used these strategies with different frequencies for emergent bilinguals after either consulting ELA or running record data (Table 36). Generally, teachers reported using instructional practices with greater frequency after viewing running record results of emergent bilingual students than after consulting ELA data of these students for 16 of these 17 instructional practices (the exception being summarizing). Therefore, the hypothesis for this analysis – teachers differ in the implementation of instructional practices as a result of either ELA or running record use – is not rejected. Furthermore, this analysis confirms the perception that teachers have of running records as more useful in the development of instructional practices for emergent bilingual students.

**Table 36. Comparison of Frequency Distribution for Types of Instructional Practices<sup>29</sup>**

Degree of Learning	ELA Grouping (%)	RR Grouping (%)	ELA Prediction (%)	RR Prediction (%)	ELA Questioning (%)	RR Questioning (%)	ELA Summarizing (%)	RR Summarizing (%)
Low	67.2	67.9	39.7	28.6	44.8	17.9	38.6	25.5
Medium	17.2	8.9	15.5	14.3	19.0	35.7	26.3	30.9
High	15.6	23.2	44.8	57.1	36.2	46.4	35.1	23.6

<sup>29</sup> Variable names are abbreviated on the table. Please refer to Table 8 for the complete list of full variable names describing each instructional practice. Instructional practices are grouped into instructional categories by color as follows: 1) grouping: gray; 2) reading strategies: light green; 3) word attack skills: light blue; 4) vocabulary: light orange; 5) background knowledge: purple; 6) home language support: light red; 7) oral language: yellow; 8) assessment: pink.

**Continuation of Table 36. Comparison of Frequency Distribution for Types of Instructional Practices**

Degree of Learning	ELA Analysis (%)	RR Analysis (%)	ELA Sight Words (%)	RR Sight Words (%)	ELA Decoding (%)	RR Decoding (%)	ELA Limited Vocabulary (%)	RR Limited Vocabulary (%)
Low	37.9	23.2	52.6	34.5	44.8	26.8	63.2	50.0
Medium	25.9	32.2	15.8	30.9	17.3	23.2	21.1	31.5
High	36.2	44.6	31.6	34.6	37.9	50.0	15.7	18.5

Degree of Learning	ELA Required Vocabulary (%)	RR Required Vocabulary (%)	ELA Unknown Words (%)	RR Unknown Words (%)	ELA Prior Knowledge (%)	RR Prior Knowledge (%)	ELA Background Knowledge (%)	RR Background Knowledge (%)
Low	37.5	20.8	33.3	14.5	37.5	16.4	36.8	16.4
Medium	33.9	26.4	24.6	21.8	12.5	18.2	12.3	21.8
High	28.6	52.8	42.1	63.7	50.0	65.4	50.9	61.8

Degree of Learning	ELA Native Language (%)	RR Native Language (%)	ELA Speak about Texts (%)	RR Speak about Texts (%)	ELA Partner Talk (%)	RR Partner Talk (%)	ELA Opportunities to Practice (%)	RR Opportunities to Practice (%)
Low	58.9	45.5	40.4	12.7	28.6	10.9	54.4	44.4
Medium	16.1	16.3	21.0	27.3	21.4	21.8	21.0	27.8
High	25.0	38.2	38.6	60.0	50.0	67.2	24.6	27.8

Degree of Learning	ELA Observe and Document (%)	RR Observe and Document (%)
Low	42.9	25.5
Medium	17.8	21.8
High	39.3	52.7

Although in comparison to the ELA data, running record data of emergent bilinguals may lead to increased use of instructional practices, it is clear that the level of implementation for most of the instructional practices after viewing running record data is not very high. In fact, only five of the seventeen instructional practices included in the survey were implemented with

high frequency –12 times or more times a month — 60% of the time after consulting running record data. These instructional practices highly implemented were:

- teaching the meaning of unknown words,
- activating prior knowledge,
- providing background knowledge,
- speaking about texts, and
- engaging in partner talk.

These instructional practices fall into three main instructional categories: vocabulary, background knowledge and oral language practices. These findings are hopeful considering that these instructional categories match what is generally thought to be best practices in reading for emergent bilinguals.

However, these results are also grim because the data does not indicate that teachers implement other instructional practices with enough frequency to truly impact the learning of emergent bilingual students. For example, the use of home language, a practice which is considered to be effective for emergent bilinguals, was infrequently used by teachers after viewing running record data. In addition, instructional practices which were grouped under the instructional category of “reading strategies” were not implemented with high frequency as a result of viewing running record data. This finding is sobering considering that it is of vital importance that emergent bilinguals’ reading development is enhanced not only by acquiring vocabulary and background knowledge, but also by developing critical comprehension skills in order to facilitate sophisticated reading skills.

Furthermore, this data was analyzed for the greatest differences between the reported frequency of use of each instructional practice between the ELA and running record data. This was done by finding the difference between high degrees of learning between ELA and running record data for each instructional practice. Although all those differences are not provided here, they are listed for the ones detailed below. The greatest differences in percentage use were found for the following instructional practices:

- teaching of required vocabulary (24.5% difference between running record and ELA high degree of learning),
- teaching meaning of unknown words (21.6% difference between running record and ELA high degree of learning),
- engaging students in speaking about texts (21.4% difference between running record and ELA high degree of learning), and

Again, these instructional practices fall under two main instructional categories: vocabulary and oral language. These results deepen the understanding of the findings reported above. Not only do teachers implement these instructional practices with greater frequency after viewing emergent bilingual running record data, but the frequency with which they implement these practices is substantially different from when they analyze ELA data. These findings signal that what teachers learn about students through running records directs them to gear up for both more vocabulary and oral language practice with their emergent bilinguals. When these findings are considered alongside the qualitative findings, the reasons behind the use of these instructional practices may be more clear. From the interviews, the case was made that although teachers report learning about their emergent bilingual students, what they learn about them is not specific to bilingualism. As such, they are acquiring student knowledge that is generic. Perhaps as a

result, teachers implement most frequently instructional practices that focused on aiding the emergent bilingual students at the initial stages of acquisition of their second language. These instructional practices, at the 3<sup>rd</sup> to 5<sup>th</sup> grade levels, would be focused on vocabulary and oral language, as was found through the analysis of this data. I would like to conjecture that if teachers acquired more bilingual-specific student knowledge about emergent bilinguals' reading development, perhaps they would implement other reading practices that represented more critical reading practices. Following these analyses, the hypothesis under consideration – teachers differ in the implementation of instructional practices as a result of either ELA or running record use – is not rejected.

The following section explores the relationship between teacher knowledge and instructional practices. This analysis consists of correlations between these two variables.

### **How are Changes in Instructional Practices Associated with Teacher Knowledge?**

The question of how reported changes in instructional practices shift when examined alongside teachers' reports of the levels of teacher knowledge acquired was explored through a correlational analysis. The hypothesis tested was:

1. Teachers' use of instructional practices is positively associated with greater levels of teacher knowledge acquired through data analysis.

Correlations between the four types of teacher knowledge – pedagogical content, student, assessment and bilingual – and instructional practices were conducted.

It is important to remember that the variables for teacher knowledge and instructional practices are both conditioned on an assessment. For example, the question to gather information about teacher knowledge was worded as, “when you look at your ELL students’

results of the New York State English Language Arts Exam, how much do you learn about?” Choices for each of the four types of teacher knowledge followed. In addition, the question which collected information about instructional practices was similarly worded. It asked respondents, “How often do you use the following instructional practices as a result of looking at the New York State English Language Arts Exam?” Therefore, there are two variables for pedagogical content knowledge, one for this type of knowledge acquired through the ELA and one for knowledge which was acquired through running records. The same is true for all types of knowledge and all instructional practices featured in this study.

The first set of correlations presented is those of knowledge acquired through the ELA with instructional practices. Table 37 displays those results. Moderate to low, significant correlations between all types of teacher knowledge acquired through the ELA and instructional practices were found.

According to these results, those teachers who reported higher degrees of acquisition of teacher knowledge from the ELA also reported using instructional practices as a result of examination of ELA data more frequently and vice-versa. Again, it is important to note that correlation analysis does not imply causality nor does not indicate a direction of the relationship.

**Table 37. Correlations between Teacher Knowledge and Instructional Practices (ELA)**

Instructional Practices	ELA Pedagogical content knowledge	ELA Student Knowledge	ELA Assessment Knowledge	ELA Bilingual Knowledge
Place Students into Reading Groups (55)	.378***	.389***	.336*	.370**
Ask Students to Make Predictions (55)	.484***	.546***	.550***	.373**
Ask Students to Generate Questions (55)	.613***	.486***	.599***	.369**
Ask Students to Summarize Texts (55)	.431***	.408***	.486***	.354**
Provide Students with Opportunities to Analyze Text (55)	.498***	.495***	.642***	.389**
Teach Students Sight Words (54)	.469***	.338*	.359**	.303*
Teach Students Decoding Strategies (55)	.501***	.543***	.499***	.352**
Provide Students with Texts with Limited Vocabulary (54)	.492***	.188	.248	.300*
Teach Students Required Vocabulary (54)	.519***	.559***	.539***	.320*
Teach Students to Figure out the Meaning of Unknown Words (54)	.487***	.525***	.525***	.357**
Active Prior Knowledge (54)	.417***	.500***	.423**	.307*
Provide Students with New Background Knowledge (54)	.438***	.517***	.434***	.311*
Provide Students with Texts in their Native Language (53)	.221	.236	.301*	.311*
Ask Students to Individually Speak about what they Read (54)	.269*	.374**	.208	.350*
Engage students in Partner Talk (53)	.607***	.646***	.534***	.464**
Provide Students with Opportunities to Practice Assessment (54)	.492***	.384**	.530***	.186
Further Observe Students (53)	.350*	.402**	.400**	.404**

\*p < .05, \*\*p < .01, \*\*\*p < .001

Note: n for each correlation is within ()

Upon further examination of the results, it is evident that a large number of moderate correlations were found between *ELA student knowledge*, the ability to tease out students' level of understanding and skills in reading, and instructional practices. Many of these moderate correlations between ELA student knowledge and instructional practices fall under the following larger instructional categories: vocabulary instruction, background knowledge and partner talk. These are the same instructional categories which were found to be implemented at a high frequency for running records. This finding signals that those teachers who report to acquire high degrees of student knowledge from the ELA data of emergent bilinguals are also likely to implement these strategies with great frequency. The use of these strategies, although helpful and important to emergent bilingual students, may also suggest the use of instructional practices, as noted above, that is geared to emergent bilingual students who are at beginning stages of developing reading skills in English, and therefore would need vocabulary and background knowledge, as well as oral language support.

In terms of other associations between the remaining types of knowledge and instructional practices, it is noteworthy that ELA assessment knowledge, or how assessments can lead to practices which support students, was strongly associated with the reading strategy of analyzing text. Perhaps this finding is due to the nature of the ELA which asks that students read passages and analyze them in order to select responses. Teachers may respond to this knowledge of what is required and respond by stepping up their use of this strategy with their students.

As for bilingual knowledge, or the understanding of how two languages develop simultaneously, acquired through the ELA, although it was positively associated with almost all instructional practices, these correlations were weak. The exception was between ELA bilingual

knowledge and engaging students in partner talk, an example of using oral language as instructional practice. This finding lends further evidence that bilingual knowledge is not driving the use and implementation of instructional practices.

In contrast to the above results, when correlations were run between teacher knowledge acquired through running records and instructional practices, almost no significant relationships were found. Those that were identified appear in Table 38 and are weak.

**Table 38. Correlations between Teacher Knowledge and Instructional Practices (RR)**

Instructional Practices	RR Pedagogical content knowledge	RR Student Knowledge	RR Assessment Knowledge	RR Bilingual Knowledge
Place Students into Reading <i>Groups</i> (55)	.136	.104	.116	.102
Ask Students to <i>Make Predictions</i> (55)	.064	.173	.209	.132
Ask Students to <i>Generate Questions</i> (55)	-.063	.201	.149	.146
Ask Students to <i>Summarize Texts</i> (54)	.099	.124	.050	.063
Provide Students with <i>Opportunities to Analyze Text</i> (55)	-.065	.175	.162	.086
Teach Students <i>Sight Words</i> (54)	.221	-.054	.038	-.062
Teach Students <i>Decoding Strategies</i> (55)	.155	.123	.203	.114
Provide Students with <i>Texts with Limited Vocabulary</i> (53)	.077	-.175	-.297*	-.136
Teach Students <i>Required Vocabulary</i> (52)	-.008	.217	-.0319	.184
Teach Students to <i>Figure out the Meaning of Unknown Words</i> (54)	.033	.217	.089	.178
<i>Active Prior Knowledge</i> (54)	.135	.254	.198	.244
Provide Students with <i>New Background Knowledge</i> (54)	.181	.275*	.275	.196
Provide Students with <i>Texts in their Native Language</i> (54)	.122	.121	.193	.231
Ask Students to <i>Individually Speak about what they Read</i> (54)	.247	.288	.207	.262
Engage students in <i>Partner Talk</i> (54)	.199	.406**	.335*	.341*
Provide Students with <i>Opportunities to Practice Assessment</i> (53)	.146	.347*	.316*	.254
<i>Further Observe Students</i> (54)	.146	.116	.016	.095

\*p < .05, \*\*p < .01, \*\*\*p < .001

Note: n for each correlation is within ()

At first the results in this table are surprising, considering that research from the literature review suggests that formative assessments directly impact student improvement. It would be assumed that improved student learning would be the result of targeted instructional practices. However, it can be explained in the findings of the qualitative data which signaled that teachers make instructional decisions *independent* of running record data, regardless of their high esteem of this assessment. These results suggest that teachers who report acquiring greater levels of knowledge from running records already implement reading instructional practices independent of the knowledge that they gain from the assessment itself. In other words, teachers employ reading instructional practices for emergent bilinguals regardless if they report learning a lot or a little from the assessment. This data lends evidence that data-driven decision making is not occurring for teachers of emergent bilinguals as a result of running record data, and that in the climate of test-based accountability, running records are being used summatively.

The contrast between the findings between instructional practices and knowledge acquired through the running records and instructional practices and knowledge acquired through ELA bolster qualitative findings. Teachers may think positively of running records, but they are not using their knowledge acquired from them to shape instructional practices. In contrast, the moderate correlations between knowledge acquired through the ELA and instructional practices may reflect the pressure that teachers voice to learn from and use these to shape instruction, although this in actuality teachers are caught within ritualized assessment practices.

In sum, the findings for correlational analysis between knowledge and instructional practices reveal that acquisition all types of knowledge through the ELA are positively associated with instructional practices. With respect to the ELA, the hypothesis – teachers' use

of instructional practices is positively associated with greater levels of teacher knowledge acquired through data analysis – cannot be rejected. However, this is not the situation between all types of knowledge acquired through running records and instructional practices. From the interviews, the case was made that knowledge, although acquired through assessments, does not impact teachers’ instructional practices. The above data analysis for running record confirms this claim and therefore the hypothesis is rejected with respect to the relationship between instructional practices and running records.

### **Summary**

The following are the findings discussed in this chapter:

1. A higher percentage of teachers report gaining greater levels of knowledge through emergent bilinguals’ running record data over ELA data. Teachers report acquiring high levels of student and bilingual knowledge as a result of examining running record data of emergent bilinguals.
2. There is a significant, positive correlation between running record assessment use and teacher knowledge acquired through running records of emergent bilinguals.
3. A higher percentage of teachers report implementing instructional practices after examining running record data over ELA data.
  - a. The instructional practices which were implemented with greatest frequency for emergent bilinguals were those related to the teaching of vocabulary, background knowledge, and oral language. Use of reading strategies and home language were used with less frequency.
4. There are significant, positive correlations between teacher knowledge acquired through the ELA and the implementation of instructional practices for emergent bilinguals. Very few exist between teacher knowledge acquired through running records and the implementation of instructional practices.

An analysis of the survey data provides a feel for teachers’ reports regarding how often they use ELA and running record data of emergent bilinguals, how much knowledge they acquire from either type of data, and the frequency with which they implement instructional

practices as a result of analyzing data. For the most part, this data both confirms what was found through analysis of the interviews and fleshes out in more detail how frequently teachers implement the given set of instructional practices.

As was found through both qualitative and quantitative analyses, for the most part teachers do not report acquiring practically-relevant amounts of teacher knowledge from either type of assessment. Furthermore, teachers report implementing some instructional practices more often after using the running record data of emergent bilinguals, as compared to when they consult ELA data of these students. These instructional practices, while worthy, reflect practices that would be most used with emergent bilinguals who are beginning to acquire reading skills in their second language rather than critical comprehension skills. Lastly, correlational analysis confirms that teachers who use running record data of emergent bilingual students make instructional decisions *independently* of the knowledge that they acquire from the data. The interpretation of the findings between knowledge acquired through the ELA and instructional practices was not as clear. Teachers in the interviews clearly stated that they were expected to learn from the ELA data and base their instructional decisions on this data. I conjecture because of this overarching pressure for teachers to report both learning and fashioning instructional practices as a result of ELA based on test-based accountability these correlations were found.

## CHAPTER 6:

### Discussion & Conclusion

Do teachers learn from the assessment data of emergent bilinguals? Does this knowledge have the potential to shape instruction for emergent bilingual students? These questions drive this doctoral thesis. The current, favored model of teaching and learning posits that instructional practices flow from assessments through the acquisition of teacher knowledge. This mixed-methods study examined if this desired course for assessment holds true for emergent bilingual students. These questions are not only academic but are also fully grounded in the political terrain of present educational practices. NCLB, the federal legislation that ushered in test-based accountability, has set the tone for future federal legislation such as Race to Top which continues to forward test-based accountability measures. An enormous amount of training, funding, and time, that appears to be ever growing, upholds the implementation of an assessment-driven agenda to improve teaching and learning in schools nationwide. Assessment-based reform has wide support, particularly among politicians and policy-makers at all levels of education administration. Though support for this type of reform is solid, the evidence whether an assessment-based movement to improve teaching and learning is not as certain, in particular for those students who have assessment needs that differ from the general population.

This dissertation research provides a sobering perspective on the current use of assessment data, both summative and formative, for emergent bilingual students. Instead of being a source of knowledge for teachers of these students, teachers are caught in a set of ritualized assessment practices that limit their abilities to glean potentially valuable information about their emergent bilingual students, thereby bypassing important opportunities to learn about

how bilingual students develop reading skills in two languages. Furthermore, it was found that the ways that summative and formative assessments are currently used do not match their intended purposes. Under the current form of test-based accountability, summative assessments are used formatively despite not being designed for this purpose. On the other hand, formative assessments, while highly regarded by teachers, are not being tapped by teachers to learn about their emergent bilingual students, nor actively used to shape instruction. Rather, teachers primarily refer to students' reading levels, thus transforming this formative assessment into a summative one in practice. Therefore, teachers are not acquiring substantial amounts of knowledge from either, nor are they implementing instructional practices as a result of the knowledge they gain from these assessments. These findings bring into question the ability of an accountability system based on the premise that assessments lead to improved and targeted instruction for emergent bilingual students to actually match its intent.

In this section, I offer an overview of general findings derived from each of the analyses. These general findings describe how assessments are, and are not, currently used by teachers as a source of teacher knowledge. In this section, I also explore the implications of these findings with respect to the following three areas:

- 1) assessment policy and practices for emergent bilinguals,
- 2) the role of teacher knowledge in catalyzing positive adjustments to assessment use in order to benefit emergent bilingual student learning, and,
- 3) the re-structuring of accountability systems to benefit emergent bilingual students.

**General Finding 1:** *Analysis of summative assessment data of emergent bilinguals does not yield useable teacher knowledge.*

*Derived from –*

*Qualitative Findings:*

1. Teachers did not report any gains in knowledge from consulting ELA data of their emergent bilingual students. They voiced a number of concerns about the validity of exam results for these students. These concerns matched those found in the literature.

*Quantitative Findings:*

2. There were no significant correlations between ELA use and types of knowledge.

Both qualitative and quantitative findings clearly demonstrate that teachers did not report acquiring knowledge from the ELA, the summative assessment studied through this research. Summative assessments are designed to reveal trends of achievement within and across student populations, not to yield information about individual student performance (Popham, 2008). Despite the intended purpose, it is common practice for both school administration and teachers to use these test results to “learn” about individual students. Furthermore, it is expected that this knowledge will drive instructional decision-making. The juxtaposition of these two opposite forces — the design of summative assessments to provide information about *groups of students* rather than individuals and the practical use of this assessment that deeply impacts individual students’ instructional opportunities — points to the confusion about the appropriate use of this assessment to drive instruction.

Educators, both administrators and teachers, are using summative assessments formatively when they were not envisioned to be used as such. Therefore, they provide limited and inaccurate information to teachers about their students. For all students, but particularly for many emergent bilingual students, using summative assessments formatively is troublesome

because these assessments are not accurate measures of these students' English reading abilities. The underlying assumption behind the use of summative assessment data is that students' scores are directly indicative of their abilities. As described earlier, for emergent bilingual students this is not the case. Therefore, any conclusions based on summative assessment results of emergent bilingual students are misguided.

Teachers are acutely aware of the limits of this assessment to evaluate their emergent bilingual students. Regardless, they are encouraged and supported by school administration to analyze the results from this exam. As described in the qualitative results, teachers *learn* how to operate within ritualized assessment practices that provides the illusion that summative data is deeply examined and leads to meaningful instructional moves. In reality, teachers simply use performance levels to identify patterns of performance and to group students despite the fact that summative assessment results of emergent bilinguals are impacted by language proficiency, as described by Abedi (2009).

The ritualized assessment practices described through teachers' narratives is shaped by a test-based accountability framework which places pressure on school leaders and teachers to demonstrate improvement in student learning through summative tests. Wherever test-based accountability is in place, teachers encounter the same pressures to use summative assessment formatively. Therefore, the claim could be made that teachers nationwide feel the same pressures to use summative data *formatively* and thus inappropriately, despite the numerous reasons why this use of data is ineffective for all students including emergent bilinguals.

The enactment of ritualized assessment practices also speaks to the lack of importance given to teacher knowledge in guiding reform to improve student learning. Ritualized

assessment practices function as operating instructions for teachers. These instructions clear the difficulty of working with the complexities of data that emerges from students who have diverse learning needs and abilities. These findings show that the way in which teachers are supported to analyze data is at odds with the acquisition of knowledge in general. In contrast to the evolving models of teacher knowledge which reflect the belief that this knowledge is critical to improving student learning, as posited by Danielson (1996) and Heritage (2009), under ritualized assessment practices the critical role of teacher knowledge in educational improvement remains undeveloped. Rather, the ways in which teachers are engaged in analyzing assessments bypasses teacher knowledge in favor of “instructions” or ritualized assessment practices, thereby “teacher-proofing” assessments. In order for teachers to acquire knowledge from the process of analyzing data of emergent bilingual students, they must exert some agency in their own learning. As demonstrated by the interviews, this was far from the case.

In test-based accountability measures, teachers’ agency and learning are not prioritized. In order for teachers to gain knowledge from emergent bilingual students’ assessments and use it to impact instruction, teachers need to direct their own learning based on the needs of their students. As detailed in qualitative findings, some teachers stated that if each item of the test was analyzed with an eye for language, then they may be able to acquire the ability to identify bias in assessment questions and perhaps even understand how to tweak questions (in their own subsequent assessments) in order to reduce bias. This is one way that teachers can direct their own learning. Furthermore, the assessment knowledge gleaned from this process could help teachers lay the groundwork for teachers’ arguments for the proper use of summative assessment data for emergent bilingual students.

Lastly, a few teachers suggested that the inclusion of emergent bilingual students into the assessment process would allow these students to direct their own learning goals. This recommendation, if implemented correctly, would allow students to demonstrate to teachers how they understand test items in terms of language and content. This would be an opportunity for teachers to gain various types of knowledge from the process, as well as help students direct their future learning. In this scenario a formative process would be linked to a summative assessment. However, if emergent bilingual students are simply provided with performance levels on the test, they would be brought meaninglessly into the process. This process would be only possible in states in which teachers receive copies of students' test booklets from the previous year.

***General Finding 2:*** *Teachers use summative assessment data primarily for the identification of levels of performance and grouping.*

*Derived from –  
Qualitative Findings:*

1. The main instructional practice that was implemented as a result of ELA assessment practices was grouping. This is potentially troublesome for emergent bilingual students who may be placed in groups that do not meet their instructional needs.

*Quantitative Findings:*

2. There are significant, positive correlations between the level of teacher knowledge acquired through the ELA and the implementation of instructional practices for emergent bilinguals.

While it was found that teachers do not acquire substantial amounts for knowledge from ELA data, the assessment data itself is “analyzed” by teachers through the ritualized practices described above. These ritualized assessment practices consist of using performance levels to sort and group students into groups so as to receive supportive services. This practice has the potential to misdirect instruction for emergent bilingual students. Emergent bilingual students

may be grouped with other students who may also have received low performance levels on the ELA, even though their scores may have been due to factors other than English language proficiency. Grouping all students who perform at low levels into one group may contribute to continued poor performance. Emergent bilingual students need specific kinds of instruction to improve their overall reading development. For example, an emergent bilingual student who has performed poorly on the ELA may need language support in English, reading instruction in Spanish, and help to transfer reading skills from Spanish to English. However, if they are grouped into a low-performing group solely through performance levels in English reading, then they are more likely to be given general help, rather than language-based help adequate for emergent bilinguals. This finding demonstrates that instruction is indeed being shaped by summative assessment results although the type of instruction that occurs has questionable meaning and impact for these students.

The finding from the quantitative analyses – moderate correlations between teacher knowledge derived from ELA analysis and instructional practices – alongside the qualitative findings is initially puzzling. An explanation for these correlations become clearer once viewed alongside the qualitative findings. Due to the pressure that teachers feel to learn from summative assessment data and subsequently to fashion instructional practices based on what they learned, there is a chance that this data reflects those aspirations rather than true learning. In other words, these correlations may reflect the ritualized assessment practices — teachers feel that they need to report that they learn from and use the assessment to inform instruction.

These findings together speak to the power of an accountability system that is unduly focused on summative assessments to spark meaningful instructional change and student

outcomes. An inappropriate use of assessments does not have the potential to change achievement of students in general and in particular for emergent bilinguals. As Menken's (2008) study of emergent bilinguals in high schools clearly demonstrates, testing deeply impacts the teaching and learning opportunities that student received. This study echoes these findings for emergent bilingual students at the elementary school level. As such, accountability models must shift to place teacher knowledge as a priority if teachers are to truly target instruction to the particularities of emergent bilingual students' needs.

One way to consider how to direct change to accountability models is to refer to the definition of "accountability." Linn (2003) describes a shared accountability system, in which all stake holders — administrators, teachers and students — are responsible for school improvement. Linn's definition is valuable for thinking about how current test-based reform can better serve the needs of emergent bilingual students. Accountability cannot be a one-size fits all proposition. Rather accountability must incorporate other criteria (other than summative assessment results) which "count" for emergent bilingual students. Teachers and administrators must go beyond the cursory use of summative assessment data. Accountability for emergent bilingual students starts with the recognition that these students are developing bilingually and the consequences this has on teaching and learning. An accountability system which is responsible to emergent bilinguals' learning needs must depart from a thorough understanding of what those needs precisely are.

**General Finding 3:** *Although teachers perceive acquiring more knowledge through the analysis of formative data of emergent bilinguals than from the analysis of summative assessment data of these students, the knowledge they do acquire is not specific to the learning needs of emergent bilinguals.*

*Derived from –*

*Qualitative Findings:*

1. Comparatively, teachers learned more from running records, a formative assessment, than the ELA, a summative assessment. Regardless, the knowledge acquired from running records was generic and reflected both a monolingual use of this assessment as well as entrenched ritualized assessment practices.
2. Teachers report acquiring student knowledge above other types of knowledge from running records. However, the student knowledge that teachers acquired was not specific to the fact that these students are emergent bilinguals.
3. Teachers did not acquire substantial levels of either assessment knowledge, pedagogical content knowledge or bilingual knowledge through the use and analysis of running records.

*Quantitative Findings:*

4. A higher percentage of teachers report implementing instructional practices after examining running record data over ELA data.
  - a. The instructional practices which were implemented with greatest frequency for emergent bilinguals were those that deal with teaching vocabulary, background knowledge and oral language. Use of reading strategies and home language were used with less frequency.

Teachers report learning a great deal from running records through the interviews and the survey. In comparison to what they reported learning from the ELA, teachers did learn more from running records. Teachers described acquiring student knowledge more than other of types of knowledge (pedagogical content, assessment, and bilingual) investigated in this research. However, the student knowledge that teachers acquired was not specific to the needs of emergent bilingual students. Furthermore, teachers did not acquire substantial amounts of

other types of knowledge required to effectively use these assessments. In sum, although teachers view formative assessment more favorably than summative ones, and do learn more from these, their potential is not realized. Formative assessments can be a platform for the development of deep understandings of individual bilingual student's reading development as well as knowledge about the teaching of reading bilingually. However, according to these results they are not presently used to their potential.

This finding has several important ramifications with respect to understanding how teacher knowledge is acquired. First, in order to catalyze the acquisition of teacher knowledge specific to emergent bilingual students it is important to understand how it develops within context. What was repeatedly found through the interview data was that the singular teachers who were able to use formative assessments effectively for emergent bilingual students were those that already possessed a certain amount of knowledge and *combined* two or more types of knowledge in order to adjust either the administration or analysis of the assessment to suit emergent bilingual students. In order to meet the needs of emergent bilingual students, it is important that teachers are able to work with combined teacher knowledge in order to adapt an assessment appropriately, analyze assessment data with an eye on bilingualism, and craft instructional practices for these students. Therefore, it is important to acknowledge that teacher knowledge is interconnected. Teachers cannot solely use assessment data to glean student knowledge as this focus does not provide the entire picture of how emergent bilingual students develop in reading. Rather, teachers must combine bilingual and student knowledge in order to develop a better grasp of these students' developing reading abilities.

Furthermore, it is critical that teachers' knowledge is *useable* (Heritage, 2007). Heritage asserts that teachers need to know how to forge the path between knowledge and practice in order for formative assessment to have a positive impact on student performance (Heritage, 2007; Heritage, 2009). Knowledge that is acquired by teachers but is not applied does not have the potential to affect student learning. I would like to push Heritage's position further. First, I would like to identify and propose a new type of knowledge – implementation knowledge (see Table 39). Implementation knowledge is the ability to transform ideas about teaching and learning and apply them to a given situation.

**Table 39. Ascenzi-Moreno's Revised Model of Knowledge for Bilingual Teachers**

<i>Bilingual Knowledge</i>	Knowledge of how two languages develop simultaneously and how to facilitate the development of language and content in two languages.
<i>Student Knowledge</i>	The ability to tease out students' "level of knowledge in a content area, understanding the concepts in the content area, the level of their skills specific to the content area, the attitudes the students are developing and their level of language proficiency" (Heritage, 2007: 142)
<i>Assessment Knowledge</i>	How assessments can lead to practices which support students.
<i>Pedagogical Content Knowledge</i>	Knowledge required for teachers to identify multiple ways to teach students within a specific topic or domain.
<i>Implementation Knowledge</i>	Knowledge required for teachers to transform ideas about teaching and learning and apply them to a given situation.

This type of knowledge is critical if teachers are expected to translate their knowledge into classroom practices that reflect the specificities of their student population. This proposed category of implementation knowledge is necessarily contextual. It is not something that can be learned academically, but rather is learned through practice, thereby necessitating teacher agency and direction. Second, in order for teachers to develop knowledge and direct their learning, the environmental conditions must facilitate and prioritize teacher learning. School administrators

must dedicate time for teacher development and knowledge construction that emerges from practice rather than being top-down (as ritualized assessment practices are). Furthermore, they must buffer teachers against the pressures of test-based accountability, so that assessments are used appropriately, and perhaps innovatively, for the benefit of emergent bilingual students.

**General Finding 4:** *Knowledge acquired through formative assessment data of emergent bilinguals was not put to use. Rather teachers' implementation of instructional practices for these students is motivated through other means.*

*Derived from –  
Qualitative Findings:*

1. Instructional practices that were implemented were loosely related to the knowledge that teachers acquired from the analysis of running record data of emergent bilinguals. Instead, instructional practices were shaped by teachers' prior store of knowledge and/or guidelines for best practice.

*Quantitative Findings:*

2. Knowledge acquired through running records was not correlated with instructional practices.

Findings showed that the analysis of running record data consisted primarily of the examination of student reading levels rather than delving into miscue analysis – which provides detailed information about student reading. Because of the weakly developed process of analyzing running records, teachers implemented reading practices based on other criteria such as their working base of knowledge acquired elsewhere or their repertoire of best practices. Both qualitative and quantitative analyses confirm this finding.

This finding that teacher knowledge from running record analysis of emergent bilinguals does not feed into instructional practices refines Black and William's work on formative assessments (1998). Their meta-analysis points to the power of formative assessments to spur on student improved learning. Although they find that formative assessments have the potential to

affect student learning in general, this research demonstrates that within the context of test-based accountability, the latent power of formative assessments is capped. Furthermore, in order for formative assessments of emergent bilingual students to catalyze student improvement, the formative assessment process must be adapted for these students.

One way in which to envision these adjustments is to refer back to the role of accommodations for emergent bilinguals with respect to summative assessments as noted in the literature review. I suggest that the idea of “accommodations” must be transferred beyond thinking of modifying the test procedure for students to being applied to teachers in analyzing assessment data, in this case formative data. As noted in the findings, most teachers do not examine formative assessment data differently for emergent bilingual students than for monolingual students. Teachers must learn to *accommodate* both the administration and analysis of assessments by taking into account bilingual language development to render these accurate measures of emergent bilingual student ability and as a valid platform for teacher learning. One way to do this is begin thinking of formative assessments more dynamically. Poehner (2007) and Gutiérrez-Clellen & Peña (2001) state that dynamic assessment have twin purposes: to assess student learning and to adapt an assessment according to student needs. I argue that in order for running records to be used to their potential for emergent bilingual students, teachers must exert agency in the ways that these are both administered and analyzed (through miscue analysis). Teacher adjustments of assessment administration and analysis must make student bilingualism central in order to adapt their use for emergent bilingual students. Moreover, teachers must also, in collaboration with other teachers and bilingualism experts, adapt miscue analysis to take into consideration bilingual students’ reading development.

**General Finding 5:** *Test-based reform as it is currently conceived and implemented does not make proper use of assessments. Teachers' analysis of assessment data, both summative and formative, is characterized by ritualized assessment practices that do not lead to the acquisition of teacher knowledge.*

*Derived from –  
Qualitative Findings:*

1. Data from the ELA was used formatively. Since the ELA is a summative assessment, it is an inadequate use of this type of data.
2. A constellation of factors including school mandates, a singular focus on student reading levels, the time required to complete running record assessments and the need for guidance in order to aid teachers to analyze these assessments with respect to language, impacted teachers' ability to gain knowledge from these assessments.
3. The almost exclusive focus on obtaining and reporting student reading levels practically rendered formative assessments summative.
4. Teachers use of assessment data was characterized by school-sponsored ritualized assessment practices that do not take into account student characteristics or teachers' knowledge about how the exam results could be made more useful.
5. Teacher understanding of student bilingualism is static and represents a model that self-enhancing rather than strengthened through the assessment-instructional cycle.

These findings hail exclusively from the qualitative portion of the study. They reveal that the process of analyzing data of emergent bilingual students or any student in general is stamped by *ritualized assessment practices* rather than being directed by teacher knowledge. This finding is sobering given that the emphasis that data-driven instruction to raise student achievement is accepted and supported in schools. In the case of summative assessments, these practices are established because they allow for the appearance that teachers are working with ELA data productively. In the case of formative assessments, assessment practices of reporting student reading levels may make the overwhelming process of working with running record data more

manageable. Regardless, ritualized assessment practices in place for both types of assessments, do not lead to solid instructional practices from assessment data of emergent bilinguals.

These results have serious implications on the power of current accountability and school improvement policies. Presently, schools are asked delve into their data in order to set out a course to improve their students' learning. However, given the findings that teachers operate within these ritualized assessment practices which do not result in positive gains of teacher knowledge or instructional practices suited for emergent bilingual students, suggests that the use of assessments may not contribute to improvement in student learning at all.

This finding stands in contrast to the original intent of including emergent bilingual students in accountability schemes – to raise these students' achievement and to raise standards for them. Unfortunately, current assessment practices do not have salubrious effects for emergent bilinguals. While teachers are locked within ritualized assessment practices, they do not acquire knowledge that is useful to improving instruction. Findings show that teachers hold misunderstandings of how to effectively teach reading to their students bilingually. Rather, they operate on ideas, such as transfer between languages, to undergird their beliefs that both language development and bilingual reading abilities will happen on its own. These beliefs create teaching “blind spots” that leave important instructional practices and opportunities for emergent bilinguals unrealized.

Therefore in order to reclaim the original intent of accountability, teachers of emergent bilingual students must not be by-passed; instead, they must be made partners in the shaping how assessments are to be used appropriately for emergent bilingual students. In this process, I contend that teachers and school leaders must exert agency in defining appropriate uses of

assessment and methods in which to analyze assessments which result in meaningful and appropriate results for emergent bilingual students.

### **Conclusion & Implications**

This research provides evidence that data-driven assessment does not occur for bilingual learners. Rather, teachers are caught within *ritualized assessment practices* which frame the use of both summative and formative assessments. In fact, these ritualized practices change the character and use of these assessments – summative assessment is used formatively and formative assessment become summative. This shift misdirects teachers in effectively using data for the benefit, in this case, of emergent bilingual students. Findings demonstrate that the assessment-instructional cycle has been steered off course for emergent bilingual students. Although test-based accountability measures aim to use assessment to close the achievement gap through monitoring of progress of students and placing high standards on all students, this study provides evidence that actual use of both of these types of assessments by teachers of emergent bilinguals runs counter to this goal. It is in fact, *because* of test-based accountability measures which place pressures on school leaders and teachers to “use data” without regard for, or despite, the intended purpose of these very assessments that ritualized assessment practices are widely carried out across schools in the study.

Both types of assessments, large-scale standardized tests and running records, will continue to be administered to emergent bilingual students. It is my hope that these will be properly administered and analyzed in order to yield learning opportunities for teachers of emergent bilingual students. As a teacher and teacher educator, I know that teachers must be learners alongside their students otherwise teaching and learning will become stalled. Most

importantly, teachers of emergent bilingual students must be tireless in learning about their emergent bilingual students as well as adjusting their knowledge to ensure that it applies to these students.

This dissertation research emerged from academic and practical interests. Based on these research findings, I advocate reversing the current trend of data-driven instruction to *student-driven assessment and analysis*. If emergent bilingual students are to benefit from the accountability movement then the administration, analysis, and use of assessments must be adapted to fit the needs of emergent bilinguals. Teachers and school administrators have a pivotal role in spurring this shift. I believe that teachers must be brave and informed about how assessments currently do not capture emergent bilinguals' skills and abilities and the ways that these assessments can be adjusted, if at all, to better suit emergent bilinguals' assessment needs. These school players must "educate up" in order for the current accountability model to shift. This is not easy since the pressures to "use" assessment data, whether it is appropriate or not, are strong.

The following recommendations spring from the research findings. It is important to note that my years working in schools both in the capacity as teacher and as coach profoundly influence my perspective on the effectiveness of recommendations in catalyzing school change. When recommendations are too numerous or have an extensive span incorporating many points, school leaders and teachers may be overwhelmed by them and therefore their implementation will not be successful. Through my experience, I have come to believe that a few powerful and narrowly focused recommendations have the potential to be incorporated into agendas for school change over time. Therefore, I have been conservative in the number of recommendations I put

forth, hoping that restricting them to a few will help guide school leaders and teachers in taking their first steps to using assessment data appropriately for emergent bilingual students.

Additionally, I have attempted to limit the majority of these recommendations to focus on the assessment of emergent bilinguals. Although I did not conduct comparative research with teachers of monolingual English students and their use of data, I would conjecture that these teachers also are caught within ritualized assessment practices which limit their learning from data of students. Therefore, recommendations 1 and 3 can be applied across the entire population of students.

### ***Recommendations:***

#### ***The Use of Summative Assessments for Emergent Bilinguals***

- 1) Summative assessments, such as the ELA, are not intended to be used formatively by educators. Nor were they intended to assess the knowledge of emergent bilinguals. Rather, their purpose is to present trends across and within populations of students. As such, school administration must stop facilitating opportunities to engage teachers in a process of using summative data to “learn” about their emergent bilingual students or their students in general.
- 2) Passages and test items from summative tests can be analyzed by teachers in order to develop knowledge of how test bias affects emergent bilingual performance. This knowledge is valuable to teachers of emergent bilinguals to engage in informed advocacy against the misuse of this test data as well as to aid them in their own construction of summative assessments.

#### ***Improvement of the Administration and Use of Formative Assessments for Emergent Bilinguals***

- 3) Formative assessments must be used as such. Full use of data from running records, rather than solely use of reading levels, must be the standard in analyzing running record data of emergent bilingual students or other students. In other words, miscue analysis adapted for bilingualism or language diversity must be practiced.

- 4) The administration of formative assessments, such as running records, must be adapted to suit the characteristics of bilingual reading development. These can include adjusting the language difficulty of the questions asked, allowing students to respond in their home language, and additional questioning in order to yield more accurate information about emergent bilingual student reading.
- 5) Teachers of emergent bilinguals must use data from both English and Spanish assessment results in order to develop a unified understanding of their bilingual students' reading development.
- 6) Parity between assessment materials in English and Spanish must exist in all schools where students are evaluated in both languages.

### ***Teacher Learning & Research***

- 7) Teachers must develop on-going knowledge about how their emergent bilingual students learn to read. Furthermore, the know-how to transform this knowledge into targeted instructional practices must be prioritized. School leaders can facilitate this by sponsoring opportunities for teachers to engage in learning about bilingualism, reading, and emergent bilingual students' reading growth.
- 8) Joint research involving academic partners and teachers must investigate how to conduct miscue analysis for emergent bilingual readers. This must remain an active and open line of research that can accommodate adjustments based on particularities of student language development and proficiencies of a variety of emergent bilingual students.
- 9) On-going professional development for teachers is necessary in order for teachers to make full use of data from running records. Through professional development focused on using running records for emergent bilinguals, teachers would learn how to administer, document, and analyze these students' reading. In particular, this work would provide teachers with an overview of emergent bilinguals' reading development that goes beyond increasing levels.

Teachers are increasingly drawn into using assessments with the idea that data-driven instruction will lead to improved learning. This dissertation research serves as a reminder that the path is not so simple and takes on different directions depending on the learning needs of students. In fact, assessment practices as they currently stand do not contribute to teacher

learning nor instructional practices that suit emergent bilingual students. If the current accountability system is to change, teachers and administrators who work with emergent bilingual students and are aware of their needs must take an active role in ensuring that the match between assessments and emergent bilingual students is a favorable one for the students. This requires a shift from thinking of instruction that is directed by data to re-thinking of the entire educational process driven by student needs. Making emergent bilingual students' needs a priority when administering and analyzing assessments has the potential to reform the impact of these assessments and ultimately to aid teachers in developing on-going knowledge of emergent bilinguals students and, most importantly, to ensure that these students receive instruction that leads to meaningful learning.

## Appendix 1. Qualitative Interview Questions

This study is about reading assessments and English Language Learners.

I am going to ask you some questions about how you learn about the English Language Learners in your class from the different kinds of reading assessments you use. I will also ask you how you use the information from these assessments to help you teach.

### Interview Protocol

The first couple of questions are intended to gather background information about your reading program, your students and your beliefs about assessments.		
Question	Rationale	Response
Please describe how you teach reading at your school.	Gather descriptive information about pedagogical techniques school-wide.	
Please describe your English Language Learners (what languages they speak, how many years in your program, their background, their learning needs).	Gather descriptive information about student population.	
How do you assess your readers at the school (both large-scale and classroom-based)?	Gather descriptive information about reading assessments at school.	
What do you like about how you assess readers? Why?	Gather teacher perspectives on the effectiveness of reading assessments?	
What would you change about how you assess readers? Why?	Gather teacher perspectives on what would make assessment process stronger.	
Do you assess your English Language Learners in reading differently or the same as above? Why?	Inquire about teachers' methods for learning about emergent bilingual's reading abilities.	

What are your beliefs about the reading assessment of English Language Learners?	Gather descriptive information about beliefs.	
How do you decide which reading assessments to use? What choices do you have in implementing these assessments?	To tie use assessments to either mandates or teacher preference.	

Now I will ask you a series of questions about large-scale standardized assessments and the classroom based assessments you use. I would like to know what you learn from each of these and how you the results from either to teach English Language Learners. (Standardized Questions)

Question	Rationale	Response
What is the purpose of (either NYS English Language Arts exam or classroom-based assessments)? Why is it important?	To uncover teacher knowledge of assessment and how it is tied to pedagogical goals.	
What is assessed and why are you assessing it?	To uncover teacher knowledge of assessment.	
Do you find this assessment gives you useful information about your students?	To inquire about if teacher uses this assessment for instructional practices.	
<i>If yes, Can you give me an example of what this assessment has told you about the <i>reading abilities</i> of a student or a group of students?</i>	To uncover teacher knowledge of assessment and knowledge of students' abilities.	
<i>If yes, Can you give me an example of what this assessment has told you about the <i>language development</i> of a student or a group of students?</i>	To uncover teacher knowledge of assessment and knowledge of bilingualism.	
Given the results of the assessment, what were your next instructional steps for the student or group of students you mentioned?	To examine the relationship between assessment and practice.	
Did you provide the student/group of students with feedback based on the results of this assessment? What was it?	To examine the relationship between assessment and practice.	

These questions are about your personal experiences with assessment and your opinion for how they should be used to improve teaching for English Language Learners. (Conversational)		
What experiences have influenced your use of assessment?	To develop a personal history of assessment use.	
Which experiences with assessment have impacted your practice most?	To examine how teacher experience has impacted practice.	
Can you give me an example of an assessment that has supported your work with English Language Learners and what you did with those results?	To gather examples of when the assessment-instructional cycle works.	
Can you give me an example of an assessment that has not supported your work with English Language Learners?	To gather examples of when the assessment-instructional cycle is not realized.	
Is there anything else you would like to share with me about what you learn from assessments of your English Language Learners?	To gather personal opinion.	
Is there anything else you would like to share with me about how you use the assessments to shape instructional practices?	To gather personal opinion.	

Lastly I will give teachers the background section from the survey.

1. What type of classroom do you teach in?

Two-way immersion (dual language bilingual)

Transitional bilingual

Other (please specify): \_\_\_\_\_

2. What grade level do you teach?

3<sup>rd</sup> Grade

4<sup>th</sup> Grade

5<sup>th</sup> Grade

Other (please specify): \_\_\_\_\_

3. How many years have you worked as a teacher? Record whole years, not fractions or months. Round up to the nearest whole number and include the current school year.

Number of years

4. How many years have you worked with bilingual students? Record whole year, not fractions or months. Round up to the nearest whole number and include the current school year.

Number of years

5. What type of teaching certification do you hold? Mark (X) ALL that apply.

- Permanent
- Alternative certification (Transitional B)
- Initial or Provisional certification
- No certification

6. Do you hold Bilingual Extension?

- Yes
- No

7. What is your highest level of education?

- BA
- Masters

## Appendix 2. Survey

### Teacher Survey

#### 1. Teacher Survey on Assessments of Bilingual Students

Thank you for your time to complete this survey. This survey is for bilingual teachers, grades 3-5, who work with English Language Learners (ELLs). This survey will take approximately 10 minutes of your time.

Data from this survey will be used by a student at the CUNY Graduate Center for a doctoral thesis. Your responses are voluntary and will be anonymous.

By filling out this survey you will receive a \$5 dollar gift card to Staples. Personal information will solely be used to mail the gift card.

Thank you again for your interest!

## Teacher Survey

### 2. Use of Assessments

These questions are about how often you use assessments to shape instruction for ELLs.

**1. During the school year how many times do you look at results from the following assessments to shape instruction for ELLs?**

	0 Times	1-3 times	4-6 times	7-8 times	More than 8 times
New York State English Language Arts Exam?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The NYSESLAT?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Formal running records (DRAs, WRAPs, etc)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conference notes of student reading?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

### 3. Knowledge from Assessments

These questions are about what you learn from the assessment of ELLs.

#### 2. When you look at your ELL students' results of the New York State English Language Arts Exam, how much do you learn about?

	Not at All	A Little	A Moderate Amount	A lot	A Great Deal
Whether my teaching of reading strategies are effective for ELLs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The skills and knowledge my ELL students possess	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How to better assess my ELL students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
About my ELL students' language development (vocabulary, use of academic language)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

4.

### 3. When you look at your ELL student's results of the NYSESLAT, how much do you learn about?

	Not at All	A Little	A Moderate Amount	A lot	A Great Deal
Whether my reading strategies are effective for ELLs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The skills and knowledge my ELL students possess	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How to better assess my ELL students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
About my ELL students' language development (vocabulary, use of academic language)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

5.

### 4. When you look at your ELL students' formal running records (DRAs, WRAPs, etc), how much do you learn about?

	Not at All	A Little	A Moderate Amount	A lot	A Great Deal
Whether my teaching of reading strategies are effective for ELLs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The skills and knowledge my ELL students possess	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How to better assess my ELL students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
About my ELL students' language development (vocabulary, use of academic language)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

6.

### 5. When you look at your ELL students' reading conference notes, how much do you learn about?

	Not at All	A Little	A Moderate Amount	A lot	A Great Deal
Whether my teaching of reading strategies are effective for ELLs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The skills and knowledge my ELL students possess	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How to better assess my ELL students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
About my ELL students' language development (vocabulary, use of academic language)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

### 7. Instructional Practices

**6. This is the first question out of three that asks you about the New York State Language Arts Exam.**

**How often do you use the following instructional strategies for ELLs as a result of looking at the New York State Language Arts Exam?**

	This Assessment Does not Impact my Use of this Strategy	1-3 times a month	1-2 times per week	3-4 times per week	Every Day
1. Regroup students into guided reading groups based on assessment data.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Ask students to make predictions prior to reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Ask students to generate their own questions about texts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Ask students to summarize texts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Provide student with opportunities to analyze texts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Teach students sight words.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Teach students decoding strategies.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

8.

**7. This is the second question that asks you about how you use the New York State English Language Arts Exam:**

**How often do you use the following instructional strategies for ELLs as a result of looking at the New York State English Language Arts Exam?**

	This Assessment				
	Does not Impact my Use of this Strategy	1-3 times a month	1-2 times per week	3-4 times per week	Every Day
8. Provide students with texts with limited vocabulary.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Teach students required vocabulary for reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Teach students how to figure out the meaning of unknown words (ex. synonyms, context clues and cognates).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. Activate prior knowledge before reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. Provide students with new background knowledge and experiences.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

9.

**8. This is the third question that asks you about how you use the New York State English Language Arts Exam:**

**How often do you use the following instructional strategies for ELLs as a result of looking at the New York State English Language Arts Exam?**

	This Assessment Does not Impact my Use of this Strategy	1-3 times a month	1-2 times per week	3-4 times per week	Every Day
13. Provide students with texts in their native language.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. Ask students to individually speak about what they read.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Engage students in partner talk about their reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Provide students with more opportunities to practice this assessment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. Further observe and document students' reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

10.

**9. This is the first question out of three that asks you about formal running records.**

**How often do you use the following instructional strategies for ELLs as a result of looking at formal running records?**

	This Assessment Does not Impact my Use of this Strategy	1-3 times a month	1-2 times per week	3-4 times per week	Every Day
1. Regroup students into guided reading groups based on assessment data.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Ask students to make predictions prior to reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Ask students to generate their own questions about texts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Ask students to summarize texts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Provide student with opportunities to analyze texts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Teach students sight words.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Teach students decoding strategies.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

11.

**10. This is the second question that asks about how you use formal running records:**

**How often do you use the following instructional strategies for ELLs as a result of looking at formal running records?**

	This Assessment				
	Does not Impact my Use of this Strategy	1-3 times a month	1-2 times per week	3-4 times per week	Every Day
8. Provide students with texts with limited vocabulary.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Teach students required vocabulary for reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Teach students how to figure out the meaning of unknown words (ex. synonyms, context clues & cognates).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. Activate prior knowledge before reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. Provide students with new background knowledge and experiences.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

12.

**11. This is the third question that asks about how you use formal running records:**

**How often do you use the following instructional strategies for ELLs as a result of looking at formal running records?**

	This Assessment				
	Does not Impact my Use of this Strategy	1-3 times a month	1-2 times per week	3-4 times per week	Every Day
13. Provide students with texts in their native language.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. Ask students to individually speak about what they read.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Engage students in partner talk about their reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Provide students with more opportunities to practice this assessment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. Further observe and document students' reading.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Teacher Survey

### 13. Part IV. Background

**12. Are you:**

- Female  
 Male

**13. What type of classroom do you teach in?**

- Two-way immersion (dual language)  
 Transitional bilingual  
 ESL

Other (please specify)

**\* 14. What grade level do you teach?**

- ESL  
 3rd Grade  
 4th Grade  
 5th Grade

**15. How many years have you worked as a teacher? Record whole years, not fractions or months. Round up to the nearest whole number and include the current school year.**

Number of years teaching:

**16. How many years have you worked with bilingual students? Record whole years, not fractions or months. Round up to the nearest whole number and include the current school year.**

Number of years teaching bilingual students:

**17. What type of teaching certification do you hold? Mark ALL that apply.**

- Permanent  
 Alternative certification (Transitional B)  
 Initial or Provisional certification  
 No certification

## Teacher Survey

**18. Do you hold a bilingual extension?**

Yes

No

**19. What is your highest level of education?**

BA

MA

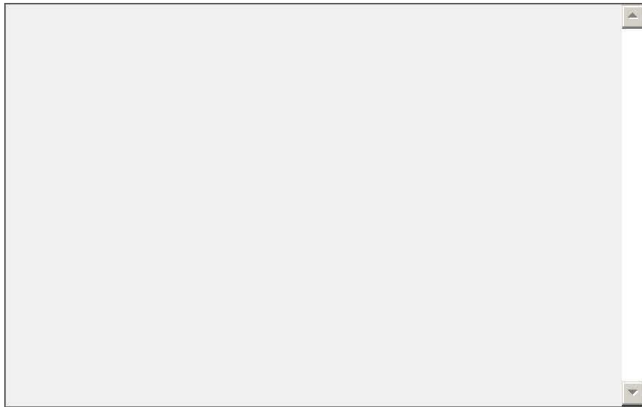
PhD

## Teacher Survey

### 14. Comments

If there is anything else you would like to tell me about how you use assessments of ELLs, the things that you learn from them and how these impact your practice, please do so here.

**20. Please write your comments here:**



### Appendix 3. Description of Quantitative Variables

Variable	Code	Value
Times ELA is Used	ELAUSE	1 = 0 times; 2 = 1-3 times; 3 = 4-6 times; 4 = 7-8 times; 5 = More than 8 times
Times NYSESLAT is Used	NYSUSE	1 = 0 times; 2 = 1-3 times; 3 = 4-6 times; 4 = 7-8 times; 5 = More than 8 times
Times Running Records are Used	RRUSE	1 = 0 times; 2 = 1-3 times; 3 = 4-6 times; 4 = 7-8 times; 5 = More than 8 times
Times Conferences are Used	CONUSE	1 = 0 times; 2 = 1-3 times; 3 = 4-6 times; 4 = 7-8 times; 5 = More than 8 times
ELA Pedagogical content knowledge	ELACONT	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
ELA Student Knowledge	ELASTUD	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
ELA Assessment Knowledge	ELAASS	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
ELA Bilingual Knowledge	ELABIL	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
NYSESLAT Pedagogical content knowledge	NYSCONT	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
NYSESLAT Student Knowledge	NYSSTUD	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
NYSESLAT Assessment Knowledge	NYSASS	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
NYSESLAT Bilingual Knowledge	NYSBIL	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
Running Records Content Know	RRCONT	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
Running Records Student Knowledge	RRSTUD	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
RR Assessment Knowledge	RRASS	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 =

		A great deal
RR Bilingual Knowledge	RRBIL	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
Conferencing Pedagogical content knowledge	CONCNT	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
Conferencing Student Knowledge	CONSTUD	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
Conferencing Assessment Knowledge	CONASS	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
Conferencing Bilingual Knowledge	CONBIL	1 = Not at all; 2 = A little; 3 = A moderate amount; 4 = A lot; 5 = A great deal
ELA Use & Grouping for Guided Reading	ELAGRP	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Prediction	ELAPRED	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Questioning	ELAQU	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Summarizing	ELASUM	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Analysis	ELAANA	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Sight Words	ELASIGHT	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Decoding	ELADECODE	1 = This assessment does not

		impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Limited Vocabulary Texts	ELALIMVOC	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Teaching of Required Vocabulary	ELAVOC	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Figuring out Meanings of Unknown Words	ELAMEAN	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Activating Prior Knowledge before Reading	ELAPRIOR	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA & Providing Students with New Background Knowledge	ELABACK	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA & Providing Native Language Texts	ELANAT	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA & Speaking about Texts	ELASPEAK	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA & Partner Talk	ELAPART	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
ELA & Assessment Practice	ELAPRAC	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2

		times a week; 4 = 3-4 times per week; 5 = Every day
ELA Use & Further Observation of Student	ELA OBS	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Grouping for Guided Reading	RRGRP	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Prediction	RRPRED	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Questioning	RRQU	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Summarizing	RRSUM	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Analysis	RRANA	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Sight Words	RRSIGHT	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Decoding	RRDECODE	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Limited Vocabulary Texts	RRLIMVOC	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per

		week; 5 = Every day
Running Record Use & Teaching of Required Vocabulary	RRVOC	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Figuring out Meanings of Unknown Words	RRMEAN	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Activating Prior Knowledge before Reading	RRPRIOR	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Providing Students with New Background Knowledge	RRBACK	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Providing Native Language Texts	RRNAT	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Speaking about Texts	RRSPEAK	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Partner Talk	RRPART	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Assessment Practice	RRPRAC	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Running Record Use & Further Observation of Student	RROBS	1 = This assessment does not impact my use of this strategy; 2 = 1-3 times a month; 3 = 1-2 times a week; 4 = 3-4 times per week; 5 = Every day
Gender	GENDER	1 = female; 2 = male

What kind of bilingual setting	SETTING	1 = Two-way immersion; 2 = Transitional bilingual; 3 = ESL; 4 = Other
Grade Taught	GRADE	1 = ESL; 2 = 3 <sup>rd</sup> Grade; 3 = 4 <sup>th</sup> Grade; 4 = 5 <sup>th</sup> Grade
How many years teaching total	TEACH	
How many years teaching bilingual students	BITEACH	
Certification	CERTIF	1 = Permanent; 2 = Alternative (Transitional B); 3 = Initial or Provisional; 4 = No certification
Bilingual Extension	BIEXT	1 = Yes; 2 = No
Level of Education	ED	1 = BA; 2 = MA; 3 = Ph.D.

**D**irections

Here is a passage from a book about comets. Read the passage. Then answer questions 1 through 7.

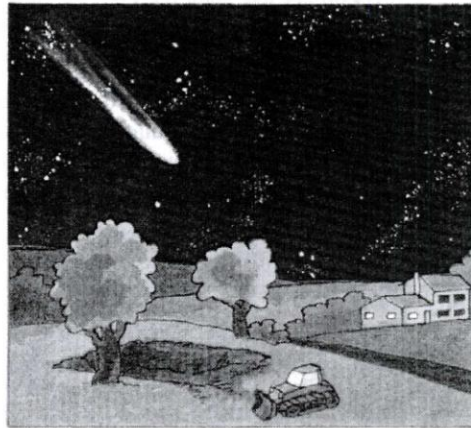
---

# Comets

by Franklyn M. Branley

illustrated by Giulio Maestro

---



Comets are parts of our solar system. Like the planets, they go around the sun.

But comets are not made of solid rock like planets. A comet is a ball of dust, stones, and ice. Many people call comets dirty snowballs.

The “snowball” may be only a few miles across. But when the sun heats the “snowball,” much of it is changed to gases. The gases expand and form the comet’s head, which may be thousands of miles across.

\* \* \*

A comet moves fast in its trip around the sun. But when we see it, the comet does not seem to move. That’s because it is so far away. The moon moves fast, but when you look at it, you can’t see any motion. That’s because the moon is far away.

So comets do not streak across the sky. You cannot see any motion. But if you look night after night, you can see that a comet changes position among the stars. If you watch the moon night after night, you’ll see that it also changes position.

\* \* \*

Each year astronomers discover new comets. Some of them are seen only once. They make one trip around the sun and then go way out into space. The sun's gravity cannot hold them.

Other comets, like Halley, keep returning. They have been captured by the sun. Halley's earliest visit was probably 3,000 years ago. It may keep returning for another 3,000 years.

But every time a comet goes around the sun, the comet loses part of itself. Gases and dust are pulled out of the comet. That's why Halley is now dimmer than it used to be. Next time it visits us, in 2062, it may be even dimmer. Each visit it may get dimmer and dimmer, until it finally disappears.

**1** What is this passage **mostly** about?

- A** comets and the sun
- B** why comets are like snowballs
- C** who discovered the first comet
- D** facts about comets

**2** According to the passage, what does a comet travel around?

- A** the sun
- B** the moon
- C** other comets
- D** other planets

**3** Why do comets look as if they are motionless when we see them?

- A** They are very small.
- B** They are very far away.
- C** They move very slowly.
- D** They move only at night.

**4** According to the passage, some comets have been "captured by the sun." What does this phrase mean?

- A** The comets are crashing into the sun.
- B** The comets have been brightened by the sun.
- C** The comets are stuck in the sun's gravity.
- D** The comets have been burned by the sun's heat.

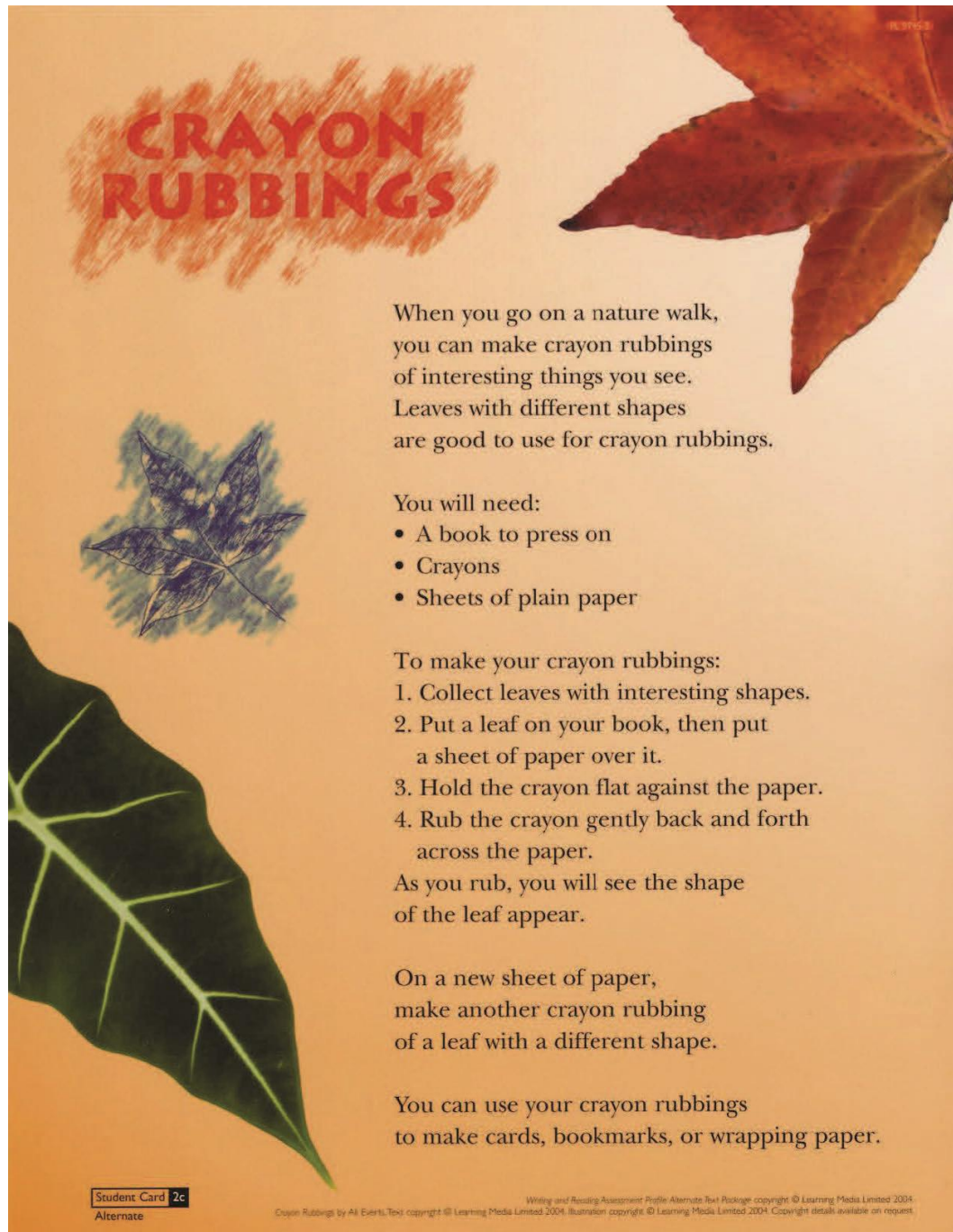
**Go On**

■ SECURE MATERIAL ■  
Do not reproduce. Do not discuss contents  
until end of designated makeup schedule.

Book 1

Page 3

## Appendix 5. Running Record Sample



# CRAYON RUBBINGS

When you go on a nature walk, you can make crayon rubbings of interesting things you see. Leaves with different shapes are good to use for crayon rubbings.

You will need:

- A book to press on
- Crayons
- Sheets of plain paper

To make your crayon rubbings:

1. Collect leaves with interesting shapes.
2. Put a leaf on your book, then put a sheet of paper over it.
3. Hold the crayon flat against the paper.
4. Rub the crayon gently back and forth across the paper.

As you rub, you will see the shape of the leaf appear.

On a new sheet of paper, make another crayon rubbing of a leaf with a different shape.

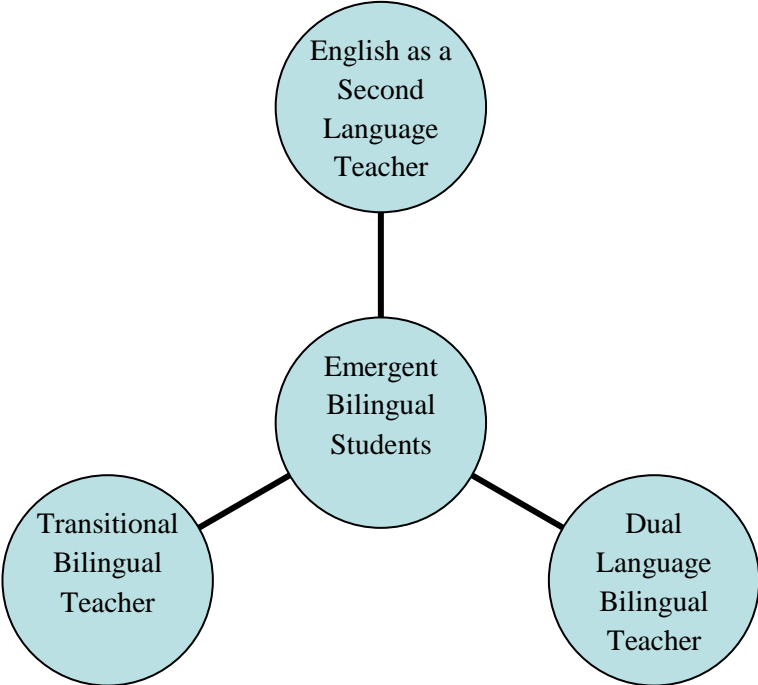
You can use your crayon rubbings to make cards, bookmarks, or wrapping paper.

**Student Card 2c**  
Alternate

Writing and Reading Assessment Profile Alternate Test Package copyright © Learning Media Limited 2004.  
Crayon Rubbings by Ali Everts, Text copyright © Learning Media Limited 2004. Illustration copyright © Learning Media Limited 2004. Copyright details available on request.



**Appendix 6. Types of Teachers that Work with Emergent Bilingual Students**



Adapted from Gottlieb (2006: 4)

## **Appendix 7. Codes for Qualitative Analysis**

### **I. ELA Knowledge**

- a. ELA Student Knowledge
- b. ELA Pedagogical content knowledge
- c. ELA Assessment Knowledge
- d. ELA Knowledge of Bilingualism

### **II. ELA Instructional Practices**

- a. Main Idea
- b. Vocabulary & Context Clues
- c. Feedback to Students
- d. Grouping
- e. Test-Taking Practices & Skills

### **III. Results of ELA Analysis**

- a. Curriculum Adaptations/Long-term Planning
- b. Evaluating Standards
- c. Planning Lessons
- d. Use as a Baseline

### **VI. Critiques of ELA vis a vis Bilingualism**

- a. Confounding Factors in the Exam
- b. English Proficiency Levels of Students

### **VII. Other Critiques of the ELA**

- a. Length of Time Tested vs. Weight of Test
- b. Difference from Reading Objectives in Class
- c. Consequences of Testing on School
- d. Teaching Testing Behaviors
- d. Differences in the Test between Grade Levels
- e. Types of Passages
- f. How the Test is Reported
- g. Imprecise Measurement

### **VIII. Running Record Knowledge**

- a. RR Student Knowledge
- b. RR Pedagogical content knowledge
- c. RR Assessment Knowledge
- d. RR Bilingual Knowledge

### **IX. Running Record Instructional Practices**

- a. Grouping
- b. Vocabulary
- c. Decoding
- d. Feedback
- e. Comprehension Strategies
- f. Identification of Story Elements
- g. Background Knowledge
- h. Language Structures
- i. Pronunciation
- j. Target Weaker Language
- k. Mini-Lessons

### **X. Running Record Analysis (How were running records analyzed?)**

- a. Oral Language
- b. Punctuation
- c. Fluency
- d. Comprehension
- e. Vocabulary
- f. Language Structures
- g. Strategies
- h. Miscue Analysis
- i. Language Proficiency
- j. Progress

### **XI. Bilingualism and Running Records**

- a. Inequities between English and Spanish Assessment Kits
- b. Bilingual Analysis
- c. Transfer
- d. Knowledge of Language

## **XII. Accommodations of Running Records for Emergent Bilinguals**

- a. For Emergent Bilinguals whose Home Language is Spanish
- b. For Emergent Bilinguals whose Home Language is English

## **XIII. Critiques of Running Records**

- a. Time
- b. Deadlines
- c. Amount
- d. Organization of Materials Required
- e. Skill Level Needed to Analyze
- f. Mandates
- g. Lack of Support to Analyze

## **XIV. Assessment Behaviors (The habits that direct how teacher implement assessments and analyze them)**

- a. RR Assessment Behaviors
- b. ELA Assessment Behaviors

## **XV. Guided Reading & Conferencing**

## **XVI. NYSESLAT**

## Appendix 8. Example of Description of Fountas & Pinnell Reading Levels

### Fountas & Pinnell Text Level Descriptions

Level	Description
<b>A</b>	<ul style="list-style-type: none"> <li>• One line of text (<i>focus on print, directionality</i>); Large spaces between words</li> <li>• Sentence structure is similar to students' language; Repeated pattern</li> <li>• Includes basic sight words</li> <li>• Punctuation includes periods, question marks, and exclamation marks</li> <li>• Pictures are highly supportive</li> <li>• Topics are familiar to children</li> <li>• Focus on a single idea</li> </ul>
<b>B</b>	<ul style="list-style-type: none"> <li>• Two lines of text (<i>return sweep</i>); Large spaces between words</li> <li>• Sentences increase in length; Sentence structure is similar to students' language</li> <li>• Repeated words or pattern</li> <li>• Includes more basic sight words</li> <li>• Includes some word endings (e.g., s, ed, ing)</li> <li>• Punctuation includes periods, question marks, exclamation marks, &amp; some commas</li> <li>• Simple dialogue</li> <li>• Pictures are highly supportive</li> <li>• Topics are familiar to children</li> <li>• Focus on a single idea</li> <li>• Setting is present, but seldom a plot</li> </ul>
<b>C</b>	<ul style="list-style-type: none"> <li>• Increased number of words and lines of text; Large spaces between words</li> <li>• Sentences increase in length and may include some embedded clauses</li> <li>• Sentence structure is similar to students' language</li> <li>• Some books have repeated words or pattern</li> <li>• Most books are about eight pages</li> <li>• Pictures are highly supportive</li> <li>• Includes more basic sight words and some compound words</li> <li>• Includes word endings (e.g., s, ed, ing)</li> <li>• Opportunities for decoding simple words</li> <li>• Punctuation includes periods, question marks, exclamation marks, and commas</li> <li>• Dialogue is frequently included</li> <li>• Topics are familiar to children, esp. experiential books [events of everyday life]</li> <li>• Characters and story plots are straightforward</li> </ul>
<b>D</b>	<ul style="list-style-type: none"> <li>• Longer, more complex stories</li> <li>• Some compound sentences conjoined by "and"</li> <li>• Simple plot but may include several elaborate episodes</li> <li>• Topics are familiar, but may include abstract or unfamiliar ideas</li> <li>• Text layout is easy to follow, but font size may vary</li> <li>• Texts range from ten to twenty pages</li> <li>• Pictures begin to extend meaning of text</li> <li>• New punctuation may be included (i.e., dashes, ellipses)</li> <li>• Larger number of high frequency words/greater variety</li> <li>• Includes more word endings, compound words, and multi-syllable words</li> <li>• More opportunities for decoding words with familiar patterns</li> </ul>
<b>E</b>	<ul style="list-style-type: none"> <li>• Sentences include more embedded phrases and clauses</li> <li>• More variety in language including some literary language</li> <li>• Topics range beyond the familiar</li> </ul>

Sources:

Fountas and Pinnell. (2001). *Leveled Books for Readers Grades 3-6*.

Fountas and Pinnell. (1999). *Matching Books to Readers: Using Leveled Books in Guided Reading*.

Novi Community School District web site. <http://www.novi.k12.mi.us/schools>

[http://www.holmdel.k12.nj.us/humanities/pdf/fountas\\_pinnell\\_level\\_descripts.pdf](http://www.holmdel.k12.nj.us/humanities/pdf/fountas_pinnell_level_descripts.pdf)

Retrieved January 9, 2012

## Appendix 9. Text Level Correlation Chart

Reading Level Correlation Chart

Grade Level	Reading Recovery	Fountas-Pinnell Guided Reading	DRA	Basal Equivilant	Lexile Levels
Kindergarten	A, B	A	A	Readiness	
	1		1		
	2	B	2	PrePrimer 1	
	3	C	3		
4	4				
Grade 1	5	D	6	PrePrimer 2	
	6				
	7	E	8	PrePrimer 3	
	8				
	9	F	10	Primer	
	10				
	11	G	12		
	12				
	13	H	14	Grade 1	200-299
	14				
	15	I	16		
	16				
Grade 2	18	J, K	20	Grade 2	300-399
	20	L, M	28		400-499
Grade 3	22	N	30	Grade 3	500-599
			34		
	24	O, P	38		600-699
Grade 4	26	Q, R, S	40	Grade 4	700-799
Grade 5	28	T, U, V	44	Grade 5	800-899
Grade 6	30	W, X, Y		Grade 6	900-999
Grade 7	32	Z		Grade 7	1000-1100
Grade 8	34	Z		Grade 8	

<http://bernardston.pioneervalley.k12.ma.us/documents/ConversionChart.pdf>

Retrieved January 9, 2012

## References

- Abedi, J., Hofstetter, C. H. & Lord, C. (Spring 2004). Assessment accommodations for English Language Learners: Implications for policy-based empirical research. *Review of Educational Research*, 74(1), 1-28.
- Abedi, J. (June 2004). Challenges in the No Child Left Behind Act for English Language Learners. *Phi Delta Kappan*, 85(10), 782-5.
- Abedi, J., Courtney, M., Leon, S., Kao, J. & Azzam, T. (November 2006). *English Language Learners and math achievement: A study of opportunity to learn and language accommodations*. Technical Report 702. National Center for Research on Evaluation, Standards (CRESST) and Student Testing, Center for the Study of Evaluation (CSE) & University of California Los Angeles (UCLA).
- Abedi, J. (2008). Classification system for English Language Learners: Issues and recommendations. *Educational Measurement: Issues and Practice*, 27(3), 17-30.
- Abedi, J. (April 14, 2009). *Examining the impact of language factors as a construct irrelevant source across ELL/non-ELL subgroups*. Unpublished manuscript. Annual Conference AERA.
- Ascenzi-Moreno, L, Espinosa, C., Ferholt, S., Loeb, M., Lugo-Salcedo, B. & Traugh, C. (May 2008). Learning through descriptive inquiry at the Cypress Hills Community School. *Language Arts*, 85(5), 392-400.
- August, D. & Hakuta, K. (eds.) (1998). *Educating language minority children*. National Academy Press. Washington, DC.
- August, D., Francis, D., Hsu, H. A. & Snow, C. E. (2006). Assessing reading comprehension in bilinguals. *The Elementary School Journal*, 107(2), 221-238.
- 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: A Co-publication of Routledge, the Center for Applied Linguistics and the International Reading Association.

- 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: 341-348.
- Bachman, L. F. (2001). *Fundamental considerations in language testing* (Fifth Ed.). Oxford: Oxford University Press.
- Baker, C. (2011). *Foundations of bilingual education and bilingualism*. Bristol: Multilingual Matters. 5<sup>th</sup> Edition.
- Bialystok, E. Majumder, S. & Martin, M.M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, 24, 27-44.
- Bialystok, E., Luk, G. & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading*, 9(1), 43-61.
- Bialystok, E. (2007). Acquisition of literacy in bilingual children: A framework for research.” *Language Learning*, 57(1): 45-47.
- Bedore, L., Peña, E., García, M. & Cortez, C. (2005). Conceptual versus monolingual scoring: When does it make a difference? *Language, speech and hearing services in schools*, 36: 188-200.
- Black, P., Harrison, C & Lee, C. (2003). *Assessment for learning: Putting it into practice*. Maidenhead, England: Open University Press.
- Black, P. & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2): 139-48.
- Butler, F. A. & Stevens, R. (2001). Standardized assessment of the content knowledge of English Language Learners K-12: Current trends and old dilemmas. *Language Testing*, 18(4), 409-427.
- Carini, P. (2000). Prospect’s Descriptive Processes. In *From another angle: Children’s strengths and school standards*. New York: Teachers College Press.

- Caesar, L. & Kohler, P. (July 2007). The state of school-based bilingual assessment: Actual practice versus recommended guidelines. *Language, Speech and Hearing Services in Schools*, 38, 190-200.
- Carlisle, J., Correnti, R., Phelps, G. & Zeng, J. (February 14, 2009). Exploration of the contribution of teachers' knowledge about reading to their student's improvement in reading. *Reading and Writing*, 22, October 5, 2009-457-486.
- Celic, C. (2009) *English Language Learners, day by day, K-6: A complete guide to literacy, content area, and language instruction*. Portsmouth, NH: Heinemann.
- Charmaz, K. (2009). Grounded theory: Objectivist and constructivist methods. In *Qualitative educational research: Readings in reflexive methodology and transformative practice*. London: Taylor & Francis Publishers.
- Chatterji, M. (2003). *Designing and using tools for educational assessment*. Boston, MA: Allyn & Bacon, Inc.
- Codó, E. (2008). Interviews and questionnaires. In *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Malden, MA: Blackwell Publishing. Eds. Li Wei and Melissa G. Meyer.
- Cochran-Smith, M. & Lytle, S. (Winter 2006). Troubling images of teaching in No Child Left Behind. *Harvard Educational Review*, 76(4), 668-693.
- Coleman, R. & Goldenberg, C. (Fall 2009). What does the research say about effective practices for English learners? *Kappa Delta Pi Record*. 10-16.
- Coltrane, B. (2002). *English Language Learners and high-stakes tests: An overview of the issues*. CAL Digests: <http://www.cal.org/resources/digest/0207coltrane.html> Retrieved February 6, 2010.
- Cooper, J., Kiger, N., Robinson, M., Slansky, J. (2010). *Literacy: Helping students construct meaning*. Belmont, CA: Wadsworth Publishing.
- Creswell, J.W. (2009). *Research Design: Qualitative, quantitative, and mixed methods approach*. Los Angeles: Sage Publishers.

- Cummins, J. (Spring 1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222-251.
- Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. & Falk, B. (November 1997). Using standards and assessments to support student learning. *Phi Delta Kappan*, 79(3), 190-199.
- Day, C., Sammons, P. & Gu. Q. Combining qualitative and quantitative methodologies in research on teacher's lives, work and effectiveness: From integration to synergy. *Educational Researcher*, 37(6), 330-342.
- Desimone, L.M. & Le Floch, K.C. (2004). Are we asking the right questions: Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, 26(1), 1-22.
- Desimone, L.M. (2006). Consider the source: Response differences among teachers, principals, and districts on survey questions about their education policy environment. *Education Policy*
- Desimone, L.M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Measurement*, 38(3), 181-199.
- Diamond, J. & Spillane, J. P. (June 2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality? *Teachers College Record*, 106(6), 1145-1176.
- Elmore, R. (2005). Accountable leadership. *The educational forum*, 69(2), 134-142.
- Escamilla, K. (2006). Semilingualism applied to the literacy behaviors of Spanish-speaking emerging bilinguals: Bi-illiteracy or emerging biliteracy? *Teachers College Record*. 108(11): 2329-2353.
- Falk, B., Ort, S. W. & Moirs, K. (2007). Keeping the focus on the child: Supporting and reporting on teaching and learning with a classroom-based performance assessment system. *Educational Assessment*, 12(1), 47-75.

- Field, A. (2009). *Discovering statistics using SPSS*. Los Angeles: Sage.
- Fowler, F. (1993). *Survey research methods*. Thousand Oaks: Sage.
- Fowler, F. (1995). *Improving survey questions: Design and evaluation*. Thousand Oaks: Sage.
- García, O. (2010). Assessments. *Forthcoming*.
- García, O. (2009). *Bilingual education in the 21<sup>st</sup> century: A global perspective*. West Sussex: Wiley-Blackwell.
- García, O., Kleifgen, J. A. & Falchi, L. (January 2008). *From English Language Learners to emergent bilinguals*.
- Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- Glass, G. & Hopkins, K. (1996). *Statistical methods in education and psychology*. Boston: Allyn & Bacon.
- Goldenberg, C. (Summer 2008). Teaching English Language Learners: What the research does – and does not – say. *American Educator*. 8-44.
- Government Accountability Office. (July 2006). *No Child left Behind Act: Assistance from education could help states better measure progress of students with limited English proficiency*. GOA-06-815
- Gottlieb, M. (2006). *Assessing English Language Learners: Bridges from language proficiency to academic achievement*. Thousand Oaks, CA: Corwin Press, 2006.
- Guitierrez-Clellen, V. & Peña, E. (October 2001). Dynamic assessment of diverse children: A tutorial. *Language, Speech and Hearing Services in Schools*, 32, 212-224.
- Gunzenhauser, M.G. & Hyde, A.M. (2007). What is the value of public school accountability? *Educational Theory*, 57(4), 489-507.
- Harvey, S. & Goudvis, A. (2007). *Strategies that work: Teaching comprehension for understanding and engagement*. Portland, Maine: Stenhouse Publishers.

- Heritage, M., Kim, J., Vendlinski, T. & Herman, J. (Fall 2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice*, 28(3), 24-31.
- Heritage, M. (October 2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, 140-145.
- Heritage, M. & Niemi, D. (2006). Toward a framework for using student mathematical representations as formative assessments. *Educational Assessment*, 11(3&4), 265-282.
- Heritage, M. & Chen, E. (May 2005). Why data skills matter in school improvement? *Phi Delta Kappan*, 86(9), 707-710.
- Harper, C. & de Jong, E. (March 2009). English language teacher expertise: the elephant in the room. *Language & Education*, 23(2), 137-151.
- Hill, H., Ball, D. Schilling, S. (2008). Unpacking pedagogical content knowledge: conceptualizing and measuring teachers' topic specific knowledge of students. *Journal for Research in Mathematics Education*. 39(4), 374-400.
- Hill, H., Blunk, M., Charalambous, C., Lewis, J., Phelps, G., Sleep, L. & Ball, D. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26, 430- 511.
- Johnson, R.B. & Onwuegbuzie, A.J. (2004). Mixed methods research: A paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Judd, C., Smith, E. & Kidder, L. (1991). *Research Methods in Social Relations*. Sixth edition. Fort Worth: Holt, Reinhart & Winston, Inc. Chapters 10 & 11.
- Kieffer, M. J., Lesaux, N. K. & Snow, C. E. (2008). Promises and pitfalls: Implications of NCLB for identifying, assessing, and educating English Language Learners. In Sunderman, G. (ed.), *Holding NCLB accountable: Achieving accountability, equity, and school reform*. Thousand Oaks, CA: Corwin Press.

- Kim Wolf, M., Griffin, N., Kao, J., Chang, S. & Rivera, N. (April 16 2009). *A review of policies on the provision of accommodations for ELLs in large-scale standardized assessments*. Unpublished manuscript. Annual Conference AERA.
- Kester, E. & Peña, E. (2002). Limitations of Current Language Testing Practices for Bilinguals. ERIC Clearinghouse on Assessment & Evaluation. ED470203.
- Kopriva, R., Emick, J., Hipolito-Delgado, C. & Cameron, C. (2007). Do proper accommodation assignments make a difference? Examining the impact of improved decision making on scores for English Language Learners. *Educational Measurement: Issues and Practice*. 26(3), 11-20.
- Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Cambridge, MA: Harvard University Press.
- Lachat, M. (1999). *What policymakers and school administrators need to know about assessment reform for English Language Learners*. Northeast and Islands Regional Educational Laboratory at Brown University, Providence, RI.
- Ladd, H., F. (January 23, 2008). Rethinking the way we hold schools accountable. *Education Week*.
- Lecelle-Peterson, M. & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English Language Learners. *Harvard Educational Review*, 64(1), 55-75.
- Lesaux, N. (2011). Lesson Plans: Nonie Lesaux. Ed. <http://www.gse.harvard.edu/news-impact/2011/05/lesson-plan-nonie-lesaux/>
- Li W. & Moyer, M. (2008). *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Malden, MA: Blackwell Publishing.
- Linn, R. L. (October 2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher*, 32(7), 3-13.
- Lyons, C. & Pinnell, G. (2001). *Systems for change in literacy education: A guide to professional development*. Portsmouth, NH: Heinemann.

- Mayer, D. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, 21(1), pp. 29-45.
- Martiniello, M. (Summer 2008). Language and the performance of English Language Learners in math word problems. *Harvard Educational Review*, 98(2), 333-368.
- Maxwell, J. (2005). *Qualitative research design: An interactive approach*. 2<sup>nd</sup> Ed. Thousand Oakes: Sage Publications.
- McCutchen, D., Green, L., Abbott, R. & Sanders, E. (2009). Further evidence for teacher knowledge: Supporting struggling readers in grades three through five. *Reading and Writing*, 22, October 5, 2009-401-423.
- McSwan, J. & Pray, L. (Fall 2005). Learning English bilingually: Age of onset of exposure and rate of acquisition among English Language Learners in a bilingual education program. *Bilingual Research Journal*. 29(3): 653-730.
- Menken, K. (2000). What are the critical issues in wide-scale assessment of English Language Learners? *NCBE Issue Brief No. 6*. Washington, DC. National Clearinghouse for Bilingual Education.  
[http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content\\_storage\\_01/0000019b/80/16/de/7e.pdf](http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/16/de/7e.pdf). Retrieved February 6, 2010.
- Menken, K. (2005). *When the test is what counts: How high-stakes testing affects language policy and the education of English Language Learners in High School*. Ph.D Dissertation. Teachers College, Columbia University.
- Menken, K. (2008). *English Learners left behind: Standardized testing as language policy*. Clevedon: Multilingual Matters Ltd.
- Mishler, E. (2009). Validation in inquiry-guided research: The role of exemplars in narrative studies. In *Qualitative educational research: Readings in reflexive methodology and transformative practice*. London: Taylor & Francis Publishers.
- Mishler, E. (1986). *Research interviewing: Context and narrative*. Cambridge, MA: Harvard University Press.

- Mokhtari, K., Rosemary, C., & Edwards, P. (2007). Making instructional decisions based on data: What, How, and Why. *The Reading Teacher*. 61(4), 354-359.
- NYCDOE, Office of English Language Learners. (Summer 2008). *New York City's English Language Learners: Demographics*.
- NYCDOE, The Division of Students with Disabilities and ELLs. (Spring 2011). *The 2010-11 Demographics of New York City's English Language Learners*.  
<http://schools.nyc.gov/NR/rdonlyres/3A4AEC4C-14BD-49C4-B2E6-8EDF5D873BE4/104510/DemoRpt0430Draft1.pdf>
- New York State Testing Program. (2010). *New York State English as a Second Language Achievement Test NYSESLAT School Administrator's Manual*.
- New York State Testing Program (2010). *Guide to the grades 3-8 testing program in language arts and mathematics*.
- Pacheco, M. (2010). English-language learners' reading achievement: Dialectical relationships between policy and practices in meaning-making opportunities. *Reading Research Quarterly*. 45(3), 292-317.
- Pearson. (2010). *Developmental Reading Assessment 2, Virtual Brochure*. Retrieved April 16, 2010.  
[http://assets.pearsonschool.com/asset\\_mgr/legacy/200944/ReaBro092801DRA2sampler\\_HR\\_28341\\_1.pdf](http://assets.pearsonschool.com/asset_mgr/legacy/200944/ReaBro092801DRA2sampler_HR_28341_1.pdf)
- Picciano, A. (2004). *Educational research primer*. London: Continuum International Publishing Group.
- Poehner, M. E. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *The Modern Language Journal*, 91(iii), 323-340.
- Popham, J. (2008). *Transformative assessment*. Alexandria, VA: ASCD.
- 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., August, D., Snow, C. & Barr, C. (2010). The interdependence continuum: A perspective on the nature of Spanish-English bilingual reading comprehension. *Bilingual Research Journal*. 33(5), 5-20.
- Rea-Dickens, P. & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*. 17(2), 215-243.
- Reese, L, Goldenberg, C. & Saunders W. (2006). Variations in reading achievement among Spanish-speaking children in different language programs: Explanations and confounds. *The Elementary School Journal*. 106(4): 363-385.
- Rivera, C. & Stansfield, C. (2001). Test accommodations for LEP students. *ERIC Clearinghouse on Assessment & Evaluation*. ED458289.
- Sato, M., Chung Wei, R. & Darling-Hammond, L. (September 2008). Improving teachers' assessment practices through professional development: The case of National Board Certification. *American Educational Research Journal*, 45(3), 669-700.
- Shulman, L. (February 1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Slavin R. & Cheung, Alan. (Summer 2005). A synthesis of research on language of reading instruction for English Language Learners. *Review of Educational Research*. 75(2): 247-284.
- Snow, C. (2008). Cross cutting themes and future research directions. In *Developing reading and writing in second-language learners: Lessons from the report of the National Literacy Panel on language-minority children and youth*. Eds. Diane August & Timothy Shanahan. New York: Routledge.
- Solano-Flores, G. (May 2008). Who is given tests in what language by whom, when and where?: The need for probabilistic views of language in the testing of English Language Learners. *Educational Researcher*, 37(4), 189-199.
- Solano-Flores, G., & Trumbull, E. (March 2003). Examining language in context: The need for new research and practice paradigms in the testing of English –language learners. *Educational Researcher*, 32(2), 3-13.

- Study of Instructional Improvement. (2003). *Teacher questionnaire – year 4*. Ann Arbor: University of Michigan. Retrieved April 16, 2010  
<http://www.sii.soe.umich.edu/documents/surveys/Teacher%20Questionnaire%202003-2004.pdf>.
- Tobias, R. (February 2009). *Principles of large-scale standardized tests*. Unpublished manuscript.
- Valdés, G. & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Westport, CT: Ablex Publishing.
- Valdés, G. (1997). Dual-language immersion programs: A cautionary note concerning the education of language-minority students. *Harvard Educational Review*. 67(3), 391-429.
- Working Group on ELL Policy. (March 26, 2010) *Improving educational outcomes for English Language Learners: Recommendations for the reauthorization of the Elementary and Secondary Education Act*. 1-13.
- Wang, L., Beckett, G. & Brown, L. (2006). Controversies of standardized assessment in school accountability reform: A critical synthesis of multidisciplinary research evidence. *Applied Measurement in Education*. 19(4), 305-328.