Assessment and Accountability: An Exploration of Teachers' Practices in Assessing English Language Proficiency

Mari B. Rasmussen

Follow this and additional works at: https://commons.und.edu/theses

Part of the Psychology Commons
ASSESSMENT AND ACCOUNTABILITY: AN EXPLORATION OF TEACHERS’ PRACTICES IN ASSESSING ENGLISH LANGUAGE PROFICIENCY

by

Mari B. Rasmussen
Bachelor of Science, University of Montana, 1980
Master of Education, University of North Dakota, 1996

A Dissertation
Submitted to the Graduate Faculty
of the
University of North Dakota
In partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Grand Forks, North Dakota
December
2008
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
This dissertation, submitted by Mari Beth Rasmussen in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

Chairperson

This dissertation meets the standards for appearance, conforms to the style and format requirements of the Graduate School of the University of North Dakota, and is hereby approved.

Dean of the Graduate School

Date
PERMISSION

Title: Assessment and Accountability: An Exploration of Teachers' Practices in Assessing English Language Learners

Department: Teaching and Learning

Degree: Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in her absence, by the chairperson of the department or the dean of the Graduate School. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Signature [signature]

Date [MM.DD.YYYY]
# TABLE OF CONTENTS

LIST OF FIGURES ......................................................................................................... x  
LIST OF TABLES ........................................................................................................ xi  
ACKNOWLEDGMENTS ............................................................................................ xiii  
ABSTRACT .................................................................................................................... xiv  

CHAPTER  
I. INTRODUCTION ..................................................................................... 1  
  Statement of Problem ..................................................................... 2  
  Purpose of Study ............................................................................ 5  
  Background/Overview ................................................................... 6  
  Research Questions ....................................................................... 16  
  Significance of the Study ............................................................. 17  
  Assumptions and Delimitations ................................................... 19  
  Definition of Terms ...................................................................... 20  
    Learner Related Terms .................................................... 21  
    Instruction Related Terms ............................................... 25  
    Assessment Related Terms.............................................. 26  
    Accountability ........................................................................ 29  
    English Language Proficiency......................................... 29  
    Terms Related to Federal Legislation.............................. 30  
  Summary ....................................................................................... 31
Organization of Study ..................................................................31

II. LITERATURE REVIEW AND THEORETICAL FRAMEWORK ......33

Introduction ...................................................................................33

Learning and Language Theory ...................................................37

Learning Theory ...........................................................................37

First Language Acquisition .............................................39

Second Language Acquisition and Proficiency .....................42

Educational Assessment ..............................................................51

Background ......................................................................51

General Education Assessment........................................53

Assessment and Theory ...................................................60

English Language Proficiency Assessment .......................62

Impact of Culture on English Language Proficiency Assessment ..............................................................65

Background and Historical Context of Educational Policies for ELLs ..............................................................67

Current Legislative Requirements ...................................76

Funding and ELL Programs .............................................79

Standards Based Education ..............................................80

Assessment and Accountability .......................................82

Current Practices in English Language Proficiency Assessment ..............................................................84

Related Research ..........................................................................89

Research Related to English Language Learners and Assessment ..............................................................89

vi
Research Related to Assessment and Teacher Impact.... 93

Summary.......................................................................................96

III. PROCEDURES......................................................................................... 98

Research Questions.....................................................................98

Design of Study.............................................................................99

Location ......................................................................................106

Participants..................................................................................108

Participants from State I ................................................113

Participants from State II ...............................................114

Participants from State III ..............................................115

Protecting Confidentiality .........................................................117

Data Collection Methods ...........................................................117

Questionnaire..............................................................................119

Interviews....................................................................................122

First Interviews...........................................................................123

Second Interviews ....................................................................125

Focus Groups ..............................................................................128

Data Analysis Procedures ..........................................................130

Questionnaire Data ....................................................................131

Interview and Focus Group Data..................................131

Coding ............................................................................133

Role of Researcher .....................................................................137

Methods for Verification and Trustworthiness .......................139

vii
Limitations ................................................................. 141
Summary .............................................................................. 142

IV. FINDINGS ............................................................................. 143

Section 1: Overview of Data From Each Activity .......... 144

Questionnaire Data .......................................................... 145
First Interview Data ........................................................ 149
Focus Group Data ............................................................ 152
Second Interview Data ..................................................... 154

Section 2: Themes .................................................................. 155

Gaps in Findings ............................................................... 156
Overview of Themes .......................................................... 158

Theme 1: Teachers use a Variety of Assessments to Understand English Language Learners ................. 160
Theme 2: Teachers Know What Their Students Can and Cannot Do ......................................................... 182
Theme 3: Teachers Bring a Variety of Backgrounds and Contexts to the Assessment Process When Looking at English Language Learners ..................................................... 188
Theme 4: The Relationship Between Teacher Developed Assessment Information and Large Scale Assessment and Accountability is Multifaceted and Mixed ......................... 196

Summary .............................................................................. 207

V. ASSERTIONS AND CONCLUSIONS ................................ 208

Introduction ........................................................................... 208
Assertions .............................................................................. 213
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Gottlieb View of Academic Language Proficiency</td>
<td>47</td>
</tr>
<tr>
<td>2.</td>
<td>The Goal of Title I and Title III Requirements: English Language Proficiency and Academic Achievement for LEP Students</td>
<td>78</td>
</tr>
<tr>
<td>3.</td>
<td>Research Design</td>
<td>101</td>
</tr>
<tr>
<td>4.</td>
<td>Example of Coding With Ethnography</td>
<td>135</td>
</tr>
<tr>
<td>5.</td>
<td>Example of Code Book</td>
<td>136</td>
</tr>
<tr>
<td>6.</td>
<td>Example of Family Tree</td>
<td>136</td>
</tr>
<tr>
<td>7.</td>
<td>Research Themes and Findings</td>
<td>159</td>
</tr>
<tr>
<td>8.</td>
<td>Research Assertions</td>
<td>215</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Standards, Assessments and Annual Measurable Objective Requirements in NCLB</td>
<td>77</td>
</tr>
<tr>
<td>2.</td>
<td>Participating States</td>
<td>107</td>
</tr>
<tr>
<td>3.</td>
<td>Research Participants</td>
<td>110</td>
</tr>
<tr>
<td>4.</td>
<td>Data Collection Activities Schedule</td>
<td>112</td>
</tr>
<tr>
<td>5.</td>
<td>Fall Interview Schedule</td>
<td>126</td>
</tr>
<tr>
<td>6.</td>
<td>Data Accounting Table</td>
<td>145</td>
</tr>
<tr>
<td>7.</td>
<td>Use of English Language Proficiency Assessment: Modification</td>
<td>146</td>
</tr>
<tr>
<td>8.</td>
<td>Use of English Language Proficiency Assessment: Student Plans</td>
<td>147</td>
</tr>
<tr>
<td>9.</td>
<td>Student Scores</td>
<td>147</td>
</tr>
<tr>
<td>10.</td>
<td>Student Grades</td>
<td>148</td>
</tr>
<tr>
<td>11.</td>
<td>State Assessment Development</td>
<td>148</td>
</tr>
<tr>
<td>12.</td>
<td>Desegregation of Student Scores</td>
<td>149</td>
</tr>
<tr>
<td>13.</td>
<td>Assessments Used by Participants</td>
<td>161</td>
</tr>
<tr>
<td>14.</td>
<td>Helpfulness of Informal Assessment</td>
<td>163</td>
</tr>
<tr>
<td>15.</td>
<td>Supplementing of Assessments</td>
<td>167</td>
</tr>
<tr>
<td>16.</td>
<td>Abstract Concepts</td>
<td>175</td>
</tr>
<tr>
<td>17.</td>
<td>Home Language Usage</td>
<td>175</td>
</tr>
<tr>
<td>18.</td>
<td>Proficiency Scoring</td>
<td>176</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

First of all, I wish to thank my advisor, Shelby Barrentine, for her countless hours assisting me with my research project. Without her guidance and direction I could not have completed this immense task. Thanks, also goes to my committee members for their support and assistance. Along with Dr. Barrentine, my committee members include Dr. Xiaozhao Huang, Dr. Nadine Tepper and Dr. Anne Walker. My family must be also be included in my thanks. They have also helped in many small and large ways. Finally, sincere appreciation goes to Sandy Krom who showed patience and humor in helping me format this paper.
ABSTRACT

Assessment is an important component in an education program. Current educational policy includes large scale assessments with stringent accountability requirements. All students are expected to meet standards for achievement, including those who have historically not been successful.

Students from non-English language backgrounds are one of the populations of students who do not achieve academic success at the same rate as other students. This population of students is increasing in our nation's schools, causing this achievement gap to be of greater concern. Added to the concern is the requirement that English Language Learners (ELLs) must meet rigorous standards in both academic achievement and English language proficiency.

The assessment of ELLs is controversial for a number of reasons. Language is highly complex and assessment must reflect that complexity. English language proficiency assessment has become a "high stakes" issue as school districts are held accountable for student progress in this area.

This research study examines the assessment of ELLs from the perspective of their teachers. Based on a theoretical foundation that supports a social constructivist view of instruction and assessment, the research design involves mixed methods with a naturalistic perspective.
Ten teachers from three states participated in the study, involving interviews, a questionnaire and an online focus group. Themes emerged from the data documenting that teachers use a variety of assessment to understand ELLs, their background and the context of assessment influences their perspectives, teachers know about their students and their relationship with the large scales assessments and accountability is multifaceted and varied.

Several assertions and recommendations were developed from these findings. The passion teachers of ELLs have for their students, supports them in maintaining a commitment despite challenges. The teachers have found a way to negotiate between assessment that is more holistic and the large scale high stakes tests. Teachers need to be supported in finding a voice in accountability systems. Assessments need to encompass holistic approaches to language. More research needs to be conducted on the construct of English language proficiency.
I dedicate this paper to the teachers of English Language Learners who
deserve praise for their hard work and commitment.
CHAPTER I

INTRODUCTION

In directing the activities of the young, society determines its own future in determining that of the young. Since the young, at a given time, will at some later date compose the society of that period, the latter’s nature will largely turn upon the direction children’s activities were given at an earlier period. This cumulative movement of action toward a later result is what is meant by growth (Dewey, 1944, p. 41).

John Dewey’s statements on education have particular significance many years later as the role of assessment in children’s learning continues to be an issue of discussion and debate. With the implementation of controversial policies and accountability systems for public schools, the assessment of students in our nation’s classrooms has become a topic for newspaper headlines, along with school staff lounge conversations and research agendas.

What future is being determined for our society by the assessment systems that are being implemented and directed upon the classrooms of today? Are we developing the sort of tools and methods consistent with the philosophy and values of a free and democratic society? Are we assessing children with systems that are reflective of the complexity of children’s growth and learning? Where do teachers and classroom assessments fit into this picture? Do the individuals who are working with students on a daily basis have a voice in assessment, accountability systems and policy?

What about the students who have traditionally lagged in achievement? As numbers of students from diverse language and cultural backgrounds increase in school
districts across the nation, the need to look at their progress in meeting educational goals becomes more significant. Many students who come to school from different language backgrounds are not proficient in the English that is necessary to succeed in an English speaking society. Current policies and legislation require that these students must also be assessed annually to ensure that they are making progress and meeting standards for English proficiency.

The following study addresses the issue of the assessment of students from diverse language backgrounds. Using a qualitative, naturalistic approach, the study looks at what is happening in an area of education and what it means to the participants. The first chapter provides an introduction to the study. This chapter includes a statement of the problem, purpose of the study, along with background information. The research questions, a discussion of the significance of the study, assumptions and delimitations, and definitions of key terms are also included in this chapter.

Statement of Problem

The assessment of students from diverse language backgrounds is a challenging area in education for a number of reasons. As the number of students from diverse language backgrounds increase in our nation’s schools, there is greater concern regarding the academic achievement of this population. Historically, these students have not achieved well in school because of their lack of proficiency in English.

Current federal legislation has put more pressure on states and school districts to ensure these students who come from backgrounds other than English. These students must not only meet the same achievement goals required of all students, but they must also meet goals for making progress in English language proficiency.
Current federal policy and legislation require states to set goals for student achievement and hold school districts to the goals. All students must make progress, including students who are not proficient in English because of a different language background. These language minority students must also make progress in English.

The state assessment systems and goals for making progress and achieving proficiency are controversial for a number of reasons. Determining whether students are making progress in English proficiency is not an easy task. Language is one of the most complex areas in education to assess. Research is not conclusive on the best way to assess English language proficiency. Research is also limited because the area of English language proficiency is relatively new.

Assessment is also a complicated and controversial area of education involving policy, practice and theory. The various questions related to assessment involve the types of assessments that should be used, what the assessments should include and what the information means. Approaches differ according to theories on how children learn. Large scale assessment systems are based on a philosophy that learning is a behavior that can be quantified and measured. This behaviorist viewpoint may not be inclusive of the complexities of human language.

Along with the conflict between educational theory and large scale assessment is the inconsistency related to the individuals involved in the assessment process. Who develops and interprets the assessments? Who develops policies that determine what assessment results are used and how they are reported? What role do teachers play? Are they involved in this process? Some claim that the only valid and reliable assessment decisions can be made by those who work most closely with the students, such as
teachers. Others believe that it is not the role of teachers. The discrepancies between
theory and current testing policies, the complexity of language and other factors,
document that there are significant challenges in the development of assessment and
accountability systems for ELLs.

Along with the issues related to the development of assessment systems, the use of
assessment and accountability also can be problematic. An important issue involves
assurance that the assessment procedures, and products that are developed are used to
empower and affirm students and teachers. A concern would be that assessment is used to
categorize and label schools and students in a way that is disengaging and
disempowering. Finally, the challenge in the discussion of English language proficiency
assessment is to assure that the understandings of ELLs English language acquisition and
achievement reflect the voices of those who know the children best.

Thus, there are many problems related to the assessment of English language
proficiency. The instruction and assessment of ELLs in an academic context is a
relatively new field with limited research. Defining proficiency in terms of “academic
English” is relatively new.

Despite the lack of research, federal requirements in the area of English language
proficiency have put pressure on states and school districts to define proficiency and
develop valid and reliable assessments. The only models available for large scale
assessment are the academic achievement assessment systems. These programs tend to
be based on learning theory that is more behaviorist in nature. This type of learning
theory may be conflicting with viewpoints of language development that emphasize the
complexity of language.
It is also uncertain whether the large scale assessment systems include the perspectives of those who understand student growth and development. Teachers are important voices in the assessment of student growth and learning according to researchers. The literature on the assessment policies, policies, programs and theories is limited in the area of teacher practice and viewpoint.

Purpose of Study

The purpose of this study is to look at the assessment of language minority or English Language Learners (ELLs) from the teachers' viewpoint. Teachers' perspectives, practices and beliefs are examined in order to find out what they are doing and what it means to them. The context of the study is the current environment in education. This current environment includes the federal policy and legislation with emphasis on standards-based assessment and accountability. It also includes the increased attention to the needs of minority students, including those from different language backgrounds. Finally, the environment of the study includes the changing dynamics within this field of education, as minorities increase in our nation's schools and there is a greater need to understand those who to teach and assess students from different languages.

The research was conducted and analyzed within the context of a theoretical background that supports a child-centered, holistic, multi-dimensional approach to language development. This theoretical foundation is reflected in the literature review, choice of data collection methods, data analysis and interpretation. The study used naturalistic research methods, supporting a social constructivist point of view.
Background and Overview

Students who lack proficiency in English because of a different language in their background are called English Language Learners (ELLs). These students come from a variety of language and cultural backgrounds. They include not only new immigrants to the United States, but also people who have been here for generations, but retained their home language. ELLs also includes indigenous people, such as Native Americans, Native Hawaiians or Alaska Natives who may have varying levels of proficiency in English.

There are greater numbers of ELLs in our nation’s schools today than in the past. “Students who speak English as a non-native language live in all areas of the United States” (Peregoy & Boyle, 2008, p. 3). The National Clearinghouse for English Language Acquisition (NCELA) reports that from 1996 to 2006 English Language Learners (ELLs) in Pre-kindergarten through 12th grade classrooms increased by 57.8% in ten years. In contrast, total enrollment increased by only 3.7% (National Clearinghouse for English Language Acquisition, 2006). These increased numbers are due to both increases in the student population, but also more accurate reporting and greater recognition of student needs.

States with the highest numbers of ELLs are California, Texas, Florida, New York and Illinois, but nearly all states have seen increases in recent years (National Clearinghouse for English Language Acquisition, 2006). This growth is expected to continue (Hill & Flynn, 2006).
Along with the increase in numbers is an increase in requirements to provide appropriate educational services and ensure that the students succeed. Historically these students have not achieved well in school.

National statistics show that ELLs are three times as likely as native English speakers to be low academic achievers. They are also twice as likely to be retained to repeat a grade. These statistics point to the challenges schools face in meeting their goal of enabling every student to succeed academically. This rapid growth in the ELL population has left many districts searching for teachers with experience and training necessary to work effectively with second language students. (Freeman & Freeman, 2008, p. 32)

Many states have requirements that ELLs must be appropriately identified, provided with educational services and assessed regularly to ensure that the programs are effective. Federal legislation includes requirements for assessment and services. Most controversial is the current authorization of the federal Elementary and Secondary Education Act (ESEA) – No Child Left Behind (NCLB) which requires that ELLs are assessed in content assessments, along with academic achievement assessments.

Historically ELLs have been excluded from mainstream instruction and mainstream assessment. This population has not achieved well in standardized academic achievement because of their lack of English. On the other hand, the assessments were not designed for students of different language and cultural backgrounds.

The focus of the current federal education policy is to ensure success of all students. There is an increased emphasis on the inclusion of students who have historically underachieved. Legislation requires that states not only include all students in the assessment systems, but also to report the results of the groups of students who have not achieved well in the past.
States must set goals for making progress in academic achievement and report the results of all schools in the state on their progress in meeting the goals. States must also disaggregate the achievement data of four groups of students that are considered at risk because of historical underachievement or absence from the assessment program. The progress of ELLs is one of the groups of students, along with racial minorities, students with disabilities and those that come from poverty backgrounds, that is reported separately from all students.

Hill and Flynn explain that the increase in growth and inclusionary assessment policies have caused a shift in responsibility for this population of students.

We used to think that the English as a second language (ESL) teacher would take care of everything. Perhaps we even encouraged classroom teachers to leave this kind of teaching to the specialists, must as we did with students in pullout special education programs…. But now, just as we have been to include special education students in our mainstream classrooms (without being told how to accomplish this feat), we are also facing the integration of growing numbers of ELLs. (2006, p. 3)

Others also agree that the current policies and the availability of data have focused awareness on the needs of ELLs. Wolf, et al. explain that “Ample data show large disparities in the achievement of ELL and non-ELL students and the need for heightened attention” (2008a, p. 6).

The policies have also created controversy and conflict. Rarely before have educational programs created headlines so frequently in local and national newspapers. USA Today questions the legislation with a headline stating “Test scores are up, but is it No Child?” (King, 2008, p. 7D). Referencing a study by the Center on Education Policy (2008), the USA Today reporter states that gaps between minorities and all students on
achievement tests have narrowed in the nation's schools, but it cannot be determined whether it is due to the NCLB requirements (King, p. 7D).

The educational success of ELLs is also in the news more than ever before. An editorial in the Boston Globe caused national attention by chastising the state of Massachusetts for neglecting this population of students. Rice and Lopez believe that the state's administration has offered no ideas to help ELLs and use the state achievement test results as demonstrating that these students are not achieving.

For ELL students in grade 8, an astounding 73 percent failed mathematics while 78 percent failed science. Among former ELL students, those who are officially considered to have reached English proficiency, 50 percent failed math and 54 percent failed science, rates of failure two to three times greater than for whites. (Rice & Lopez, 2008, p. 5)

All minority students in the nation are not faring well according to Little and McCarty. They express concern that the current policies and NCLB have not improved educational programs for minority students.

Recent research and a national survey have documented the unintended negative consequences of NCLB for the students, communities, and schools it most directly affects. A U. S. Civil Rights Commission report found that NCLB has done little to close the achievement gap. The report notes that the policy's prescriptive nature, its high stakes for minority students and schools, and its failure to close the gap in financial resources between the richest and poorest districts are actually widening the gap between children of color and their more affluent white peers. (Little & McCarty, 2006, p. 28)

In contrast to the news articles and research reports that question NCLB and criticize government administrations for neglecting minorities, the United States Department of Education (USDE) has disseminated information on the success of the current policies and legislation. This success is referenced on the official website for NCLB and the USDE, along with public speaking engagements by Department staff.
“Every day we learn what works so students can make greater progress,” we are told in the report, *Mapping America’s Educational Progress* (USDE, n.d.), which includes data on student increases in achievement. Secretary of Education, Margaret Spellings, applauds the progress the nation’s schools have made in setting standards and meeting educational goals for English Language Learners (ELLs) in academic achievement. “We now have proof that high standards and accountability are paying off. Our national report card shows that scores are rising and English language learners are achieving record highs” (Spellings, 2005).

Are schools and teachers successful in their goals of educating students today? Are the students from minority populations, including ELLs, making progress and learning academics as some say or are the gaps widening between the different populations of students as others say?

There are many perspectives on this subject. State and national test scores compete with the voices of researchers, teachers, and others who present different pictures of children and achievement. Each picture can be convincing. Similar to other information in our culture today, the presentation of a concept can become just as important in convincing the audience of its reality as the concept.

Anderson discusses the multiple perspectives available in our world today from a post modern perspective. “Surrounded by so many truths, we can’t help but revise our concept of truth itself; our beliefs about belief” (1995, p. 8). He continues in saying “more and more people become acquainted with the idea that, as philosopher Richard Rorty puts it, truth is made rather than found” (1995, p. 8). Sorting out student achievement in today’s educational world with the various observations and perspectives
available through the media and other sources, can cause one to wonder which reality to accept. Along with questioning which perspective is the most plausible picture of what is happening in school for students, others issues surface. How should students be assessed? Who best knows how students are doing?

An attempt to understand student assessment with a focus on language minority students involves an understanding of what assessment and the assessment of ELLs involves. Ultimately, an assessment tool or system, whether it is standardized, norm-referenced, criterion-referenced, alternative, summative or formative, needs to provide accurate and comprehensible information on student learning, progress and achievement. If judgments are to be made on students, school districts, teachers, or programs, the information provided by the assessment system must provide a picture that is defensible.

The development of English language proficiency assessment systems that are required by NCLB is challenging because of the difficulties of assessing language. Gottlieb states “Assessment of English language learners is a more complex undertaking than assessment of proficient English-speaking students because it involves the documentation of both language proficiency and academic achievement” (2006, p. 8).

All assessments are based on an understanding or theory of student learning. There are a number of theories of language learning and acquisition. Lantolf discusses some of the differences in theories in his explanation of a social constructivist approach to language learning, which supports the multi-dimensional aspect of language. “Unlike most theories of language acquisition, in particular that espoused by Chomsky’s innatist theory, the sociocultural perspective recognizes that humans are not completely at the mercy of their biology” (Lantolf, 2004, p. 349). He explains and describes how
language learning involves not only inborn, automatic responses, but also creative
eexpression, problem-solving and social interaction.

Most large scale assessment systems tend to be based on models that support
learning that are primarily behavioral or empirical in which learning products or
processes can be measured. Student progress can then be quantified in data that can be
analyzed against pre-determined formulas.

Measurement and assessment have always been important elements in assisting
teachers and administrators evaluate student learning. Education has historically
involved some sort of assessment that provided reports to parents and others. At the time
Dewey was discussing the role of growth in education, the teacher was the primary
interpreter of that growth. She or he was the expert on the child’s progress and
development.

As technology provided tools for mass production and education emphasized
consistent, quantifiable information, commercially developed, standardized assessments
have increased over the years.

Today, the sorting function of testing dominates. First used in World War I to
determine military recruits’ suitability as officers, large-scale standardized tests
sort, track, and stratify individuals and groups, separating the qualified from those
judged less qualified – for employment, higher education, and the professions.”
(Scherer, 2005, p. 9)

Current federal education legislation continues the trend towards a scientific,
behaviorist approach to assessment with increased requirements for large-scale,
standards-based state assessments to document student progress and school district
success. NCLB has more than doubled the amount of testing required of states. The
dollars spent on the development, publishing, administration, grading, and organizing of
state exams has increased from an average of $457,000 in 2004-2005 per state to a projected $570,000 in 2007-2008 (Toch, 2006, p. 8). Despite the increase in dollars spent towards testing, the question of whether these assessment systems provide a true picture of children’s progress and achievement continues to be debated.

Toch addresses some of the challenges of the large-scale state assessment programs in appropriately testing students. “Creating high-quality tests is difficult and labor intensive” (2006, p. 8) he states.

Do multiple-choice, true and false, and short answer questions truly assess knowledge? Shouldn’t tests include more “real life” type questions? But, as test developers include more items that address actual student performance, such as writing tasks, costs dramatically increase. Bubble sheets can be scored by a machine. (Toch, 2006, p. 14)

Toch (2006) finds several concerns that impact the quality of the current assessment systems required under NCLB. They include such issues as the shortage of psychometricians who are trained in measurement theory, lack of capacity of state education agencies to implement assessment programs, over-dependence on the testing industry, and demands put upon small states. The consequences of these problems ultimately result in flaws in the system, such as tests that include inaccurate scoring, items that assess only basic skills, and results that are received months later.

To further complicate the issues surrounding the validity and reliability of current assessment tools are the philosophical and political undercurrents of much of the discussion. Though there are different kinds of assessment, with different purposes, Serafini explains that “Each of these assessments may also involve different beliefs about the nature of knowledge, the level of teacher and student involvement, the criteria for
evaluating student achievement, and the effects of these assessment frameworks on classroom instruction” (2005, p. 87).

Into this cacophony of voices enters a newcomer to the world of assessment and accountability: English language proficiency assessment. As the numbers of students grow, assessment information is needed in order to develop appropriate instructional programs, ensure children are making progress, and provide information to stakeholders. “The research on the development of English literacy strongly suggests that adequate assessments are essential for gauging the individual strengths and weaknesses of language-minority students, making placement decisions, and tailoring instruction to meet student needs” (August, 2006, p. 7-8).

Several areas need to be addressed in English language proficiency assessment. It is essential for assessment to be grounded in appropriate theory and belief systems (Serafini, 2005). The construct of English language proficiency needs to be defined. “The complexity of language and the lack of consensus as to the exact nature of language proficiency is critical for one fundamental reason. Each language proficiency test should be based on a defensible model or definition of language proficiency” (Del Vecchio & Guerrero, 1996, p. 4). It is also important for teachers to be part of the discussion. “If students construct information as they learn, and apply the information in classroom settings, assessment should provide the students with opportunities to construct responses and apply their learning to problems that mirror their classroom activities in authentic ways” (O’Malley & Pierce, 1996, p. 10).

Johnston and Costello discuss literacy assessment from a social constructive viewpoint, explaining the incompatibility of this theory with the large scale assessments
currently used to assess students’ progress. They explain that “Although assessment
often is viewed as a technical matter of developing accurate measuring instruments, it is
more centrally a set of social practices in which various tools are used for various
purposes” (2005, p. 258).

In their discussion, Johnston and Costello emphasize that all assessments,
including the large scale summative tests used for accountability, along with the teacher
developed formative assessments, need to be “grounded in current and consistent
understandings of learning” (2005, p. 259). Though the learning theories of testing
psychometric practices have been implicit, they state, they tend to be individualistic and
behaviorist. “For example, current accountability testing, driven by psychometrics, is
based on rewarding and punishing students, teachers, and school systems” (Johnston &

Johnston and Costello also discuss the fact that the large scale assessments do not
always reflect information from those who best know students. In fact, the systems tend
to support structures that place less value on teacher and classroom level assessment.

Assessment discourses distribute and sustain power relationships. For example,
formative assessments, while grounded in current understandings of learning, are
not taken seriously as a form of assessment (Black & William, 1998a). They are
referred to as “informal,” as opposed to the more authoritative “formal”
assessments. There are probably many reasons for their lack of institutional
power aside from the fact that they don’t always involve a textual record or
artifact such as running records, documented events, or writing samples. They are
the purview of teachers, mostly women, and they are normally not in the language
of mathematics. When brought to a Committee on Special Education meeting,
these assessments are easily trumped by the tests of the school psychologist.
(2005, p. 263)

The important role of the teacher in an assessment system is also expressed in the
“Standards for the Assessment of Reading and Writing”, which were developed by the
International Reading Association and National Council of Teachers of English Joint Task Force. Standard 7 states “The teacher is the most important agent of assessment” (International Reading Association, 1994, p. 18).

Thus, there are many issues that complicate the field of English language proficiency assessment. These issues involve the various disconnections between theories, practice, policies and practitioners.

Research Questions

The research questions in this study reflect the complexity of the research topic and problem. They are developed to support a study that would address areas of concern and also allow for new information that may be relevant, to emerge.

The researcher must begin with the nature of the research question. According to our assumptions of science, the research question must be considered first, the accessibility of the data second, and whether the data are or are not quantified, according to the design of the study, third. In other words, the decision about what data to collect, as well as what to do with those data after they are collected, should be dictated by the research question. (Newman & Benz, 1998, p. 15)

The questions include:

1. What assessments are English as a Second Language (ESL), English Language Learner (ELL), and bilingual teachers using to document the development of English language proficiency of English Language Learners?

2. What do the teachers say the assessments reveal about the students?

3. What is the relationship between teacher developed assessment information and large-scale assessment data and accountability?

The questions focus on assessment of ELLs from the teacher perspective. They are broad enough to allow me to explore various related issues such as the assessment
tools that teachers prefer, opinions on assessing progress and exiting students, along with viewpoints on their own voice and relationship with the federal and state assessment requirements. The questions also support data collection methods that are primarily qualitative. Qualitative interviewing is most suited to collect data to address the research problem and questions.

Significance of the Study

This study has significance for many reasons. As discussed, the topic of English language proficiency assessment is controversial because of the lack of research, differences of opinion on who should be involved with assessment, what should be assessed and what the information means. The study explores an important issue and meets a need for more information. It addresses the question of whether social constructive language learner theory which emphasizes the complexity of language and the important role of teachers in assessment is reflected in current assessment tool development.

Assessment has become a challenging and critical topic in all of education. Assessment of students with limited English proficiency is particularly crucial because as this population increases in numbers, greater attention is being given to their educational needs and assessment is an integral part of a successful educational program.

"Understanding the issues concerning instruction, assessment and classification of English language learner (ELL) students is of the utmost importance given the fact that ELL students are the fastest growing student population in the United States" (Abedi, 2007, p. 3).
The fact that new assessments have been developed for this population of students becomes even more significant for many reasons. Research on English language proficiency assessment is limited or lacking (August, 2006; Abedi, 2007). The limitation of fiscal resources is also an issue that can cause state assessment programs to lack quality and depth (Toch, 2005). These assessment systems are vulnerable to the same pitfalls that have befallen the academic achievement assessment programs. These “pitfalls” includes the lack of capable assessment personnel at the state levels, politics, and power struggles impacting key decisions and the negotiations for changes and adjustments in requirements that creates systems that are more dependant on convoluted formulas than assessment of real children (Johnston & Costello, 2005; Toch, 2005).

The lack of research and information on the construct of English language proficiency is also a significant issue. More information is needed on what it means to be proficient in English for ELLs in the nation’s classrooms. What is expected for students in today’s classrooms? Bailey states:

There are major research and policy problems facing the United States and other English – speaking countries with large populations of primary and secondary school students learning English in academic contexts for academic purposes. The problems can be more succinctly articulated as a lack of comprehensive information about what language demands are placed on school-age children in general and how much English language learners can realistically be expected to learn and how quickly. While we have a large literature base about English as a second language (ESL), this research base cannot answer these questions with sufficient specificity to aid policy makers and educators faced with the creation of English language development tests and curricula. (2007, pp. 2-3)

In summary, the significance of this research is that it addresses critical and timely issues in an uncommon manner. It addresses questions that are being debated by high profile policymakers, government officials, and others in positions of authority and
power, but looks for the answers in a place that most would not consider. The study looks for truth in the hearts, voices, and thoughts of the individuals who, perhaps, have the least amount of input in the process of developing policy and procedures – the teachers.

Assumptions and Delimitations

The study includes a number of assumptions and delimitations. Assumptions can be considered facts and understandings that are basic to the research study and design, but are not necessarily verified through the methods and procedures (Roberts, 2004, p. 129). Delimitations are factors that establish the boundaries of the study and are in the control of the researcher (Roberts, 2004, pp. 128-129).

The assumptions inherent in this study relate to beliefs intrinsic to the educational philosophy and learning theory that the research is based upon. Basic to the research is the assumption that teachers have a certain amount of preparation to instruct and assess student progress. Parallel to that assumption is the expectation that schools exist to educate and prepare students for future lives. This assumption does not go further into any expectations on what teachers should know and be able to articulate about their students.

There is also an assumption that student progress in educational programs needs to be assessed. Despite different theories, philosophies and viewpoints, the literature is consistent on the expectation that instructional programs must include some sort of assessment. This assumption would not include, though, the types of assessments and other related issues.
The delimitations of this study are minimal and include the very issues that could be considered its strengths. They include:

- Nine to eleven participants,
- Setting of research location involving three states, and

The study has specific boundaries which are appropriate for the research problem, questions and design which will be discussed in Chapter III. The participation of nine to 12 participants was determined to be a sufficient number to address the issues. Three states were also considered sufficient. The time frame allowed for all the activities to be completed.

**Definition of Terms**

Both the areas of assessment and the field of ELL have many terms with definitions unique to the subject. Definitions can vary according to theory and educational philosophy. Brown explains:

A definition is a statement that captures the key features of a concept. Those features may vary, depending on your own (or the lexicographer’s) understanding of the construct. And, most important, [sic] that understanding is essentially a ‘theory’ that explicated the construct. So a definition of a term may be thought of as a condensed version of a theory. Conversely, a theory is simply – or not so simply – an extended definition. Defining, therefore, is serious business: it requires choices about which facets of something are worthy of being included. (2007, p. 5)

Key terms that are essential to the study will be defined. The definitions provided are considered operational definitions that can be used to clarify their meaning as used in the context of this study (Roberts, 2004, p. 129).
Learner Related Terms

There are many terms used in education today to refer to students from backgrounds where a language other than English is used. Whereas many are interchangeable, some have slight differences. This study concerns a category of students that have come to be popularly known as "English Language Learners". The acronym "ELL" is often used and will be used in this study. Many other terms have been used for this population of students, though. Whereas most of these terms are not used in the study, they are provided in an effort to give a context for the study. The various terms reflect the multiplicity of meanings and understandings for these students. Some of the other terms that are used for these students include:

ESL learner: English as a second language learners

ESOL learners: English as a second or other language learners, or English speakers of other languages

CLD students: Culturally/linguistically diverse students

LEP: Limited English proficient and,

PEP: Potentially English proficient.

Limited English proficient (LEP): is the term that is used in federal legislation for ELL. Because USDE uses ELL interchangeably with LEP, ELL will be used in this study with the understanding that the definition is consistent with the federal definition of LEP in legislation.

The term "limited English proficient", which is defined in section 9101 of Title IX when used with respect to an individual, means an individual -

- who is aged 3 through 21;
• who is enrolled or preparing to enroll in an elementary school or secondary school;
• who is a Native American or Alaska Native, or a native resident of the outlying areas; and
• who comes from an environment where a language other than English has had a significant impact on the individual’s level of English language proficiency; or
• who is migratory, whose native language is a language other than English, and who comes from an environment where a language other than English is dominant; and
• whose difficulties in speaking, reading, writing, or understanding the English language may be sufficient to deny the individual -
  o the ability to meet the State’s proficient level of achievement on State assessments described in section 1111(b)(3);
  o the ability to successfully achieve in classrooms where the language of instruction is English; or
  o the opportunity to participate fully in society. (USDE, 2001, Title IX, Section 9101, (25)

Federal requirements and accountability systems have put more emphasis on the classification of students and formalized many terms that had been used informally before. Once ELLs attain proficiency and exit the status of being considered ELL, they enter a classification of “former ELL” or former LEP.

FEP: Fully English proficient, and
Students who qualify as ELL are often categorized by their level of proficiency. Most states have consistent labels for the various levels that go from beginning to proficient.

Gottlieb (2006) explains that ELLs follow a series of predictable, developmental stages of language proficiency that form a continuum:

This continuum is arbitrarily divided into levels, from little proficiency to that of being proficient in English. By knowing the language proficiency level of their second language learners, teachers can better plan instruction and assessment to meet the students’ individual needs. (Gottlieb, 2006, p. 26).

The International Teachers of English to Speakers of Other Languages (TESOL) uses five levels in their English language proficiency standards (TESOL, 2006). These levels include:

Level 1-Starting
At L1, students initially have limited or no understanding of English. They rarely use English for communication. They respond nonverbally to simple commands, statements, and questions. As their oral comprehension increases, they begin to imitate the verbalizations of others by using single words or simple phrases, and they begin to use English spontaneously. At the earliest stage, these learners construct meaning from text primarily through illustrations, graphs, maps, and tables.

Level 2-Emerging
At L2, students can understand phrases and short sentences. They can communicate limited information in simple everyday and routine situations by using memorized phrases, groups of words, and formulae. They can use selected simple structures correctly but still systematically produce basic errors. Students begin to use general academic vocabulary and familiar everyday expressions. Errors in writing are present that often hinder communication.

Level 3-Developing
At L3, students understand more complex speech but still may require some repetition. They use English spontaneously but may have difficulty expressing all their thoughts due to a restricted vocabulary and a limited command of language structure. Students at this level speak in simple sentences, which are
comprehensible and appropriate, but which are frequently marked by grammatical errors. Proficiency in reading may vary considerably. Students are most successful constructing meaning from texts for which they have background knowledge upon which to build.

Level 4-Expanding
At L4, students' language skills are adequate for most day-to-day communication needs. They communicate in English in new or unfamiliar settings but have occasional difficulty with complex structures and abstract academic concepts. Students at this level may read with considerable fluency and are able to locate and identify the specific facts within the text. However, they may not understand texts in which the concepts are presented in a decontextualized manner, the sentence structure is complex, or the vocabulary is abstract or has multiple meanings. They can read independently but may have occasional comprehension problems, especially when processing grade-level information.

Level 5-Bridging
At L5, students can express themselves fluently and spontaneously on a wide range of personal, general, academic, or social topics in a variety of contexts. They are poised to function in an environment with native speaking peers with minimal language support or guidance. Students have a good command of technical and academic vocabulary as well of idiomatic expressions and colloquialisms. They can produce clear, smoothly flowing, well-structured texts of differing lengths and degrees of linguistic complexity. Errors are minimal, difficult to spot, and generally corrected when they occur. (TESOL, 2006, Standards website)

Transitioning through the different levels of English language proficiency is an individual issue. Background factors can contribute to a student's rate of progress in learning English. Within the ELL category are several types of students that are defined according to their characteristics. While these categories are not specifically addressed in this study, they related to subject since they describe some of the student populations that the participants work with. Immigrant students who come to the United States lacking in years of education or have had very different educational experiences are considered “limited formal schooling”.

24
“Long-term ELLs” is another term that is used for students who have been exposed to English for some time and may have benefited from English development services, but have not yet achieved proficiency. Freeman and Freeman state that “What is most disturbing about long-term English learners is that there are so many of them” (2008, p. 37).

Students may also be categorized according to their background. Students who come to the United States through refugee resettlement are considered “refugees” and have many unique needs because of the traumas and hardships they have experienced. “New immigrant ELLs” may be refugees, but they also may be students who have come to American schools for reasons other than refugee status. Many school districts with large numbers of new immigrant and refugee ELLs categorize these students as “newcomers” and implement “newcomer” or “welcome” educational centers which provide focused services.

ELLs can be considered part of a larger group of students that are referred to as “language minority students”. Language minority students come from another language background and may or may not be limited in English. Bilingual learners, similar to language minority students, may or may not be limited in English. The term simply identifies that they have a language other than English in their background. Bilingual learners may be fluent in their first language or they may be learning it along with English in a revival or heritage language program.

**Instruction Related Terms**

The type of instruction ELLs are provided with can fit into different categories. English language development is a term that refers to specialized instruction that supports
the student's English development. There are many different instructional programs that are used with ELLs. Traditionally the instruction and the teachers have been referred to as English as a second language. Bilingual education refers to instruction that involves another language. Typically this other language is the home language of the student. There are a number of different types of bilingual education.

Assessment Related Terms

There are many terms related to educational assessment. Most important to consider is the term “assessment” itself. Hill, Ruptic and Norwick provide a straightforward definition, stating “Assessment is the process of gathering evidence in order to document learning and growth” (1998), p. 15). A more complete definition from a document disseminated by the National Clearinghouse for English Language Acquisition addresses the issue with more detail.

Assessment is a broad term that involves the collection and maintenance of various types of data about students including norm-referenced tests, criterion-referenced tests, classroom-based assessments of various types, and performance-based tasks. We use the term “assessment” throughout this document to refer to any situation in which students must respond to items or tasks in order to demonstrate their knowledge and/or skills in a specific area. Using the appropriate type of assessment for a specific purpose is important to the validity and fairness of that assessment. A particular assessment can be reliable, valid, and fair for one purpose, but not for another. For instance, the Iowa Test of Basic Skills (ITBS) may be valid, reliable, and fair for measuring language arts achievement, not for measuring English language proficiency. (Wilde, 2004, p. 1)

There are different kinds of assessments. It is an area that is difficult to define and categorize because terms overlap and mean different things to different people. Generally speaking, assessments can be broadly divided into formal assessments and informal activities. The use of informal and formal can be controversial though, because of an inherent value judgment on something that is considered informal. The inference is
that it lacks the credibility of a formal assessment. Alternative assessment is another
term used instead of informal. This term also can be misunderstood as being second best.
I will use the terms formal and informal in this study for two reasons. The use of these
terms is consistent with the literature. I would also argue that, contrary to the implication
of lacking credibility, the term informal can reflect a more comfortable, authentic and
relaxed style, consistent with a theoretical understanding that emphasizes the creativity
and unpredictability of children's learning.

Both informal and alternative assessments include activities that are performance
based, authentic and usually involve the teacher more closely. Peregoy and Boyle
explain that “informal assessment measures include such items as teacher-made tests,
miscue analysis of oral reading, checklists, anecdotal observations, and student work
samples” 2008, p. 105).

Authentic assessment is other term that is used often in discussing alternative or
informal assessment. O'Malley and Pierce define use the term authentic assessment to
“describe the multiple forms of assessment that reflect student learning, achievement,
motivation, and attitudes on instructionally-relevant classroom activities. Examples of
authentic assessment include performance assessment, portfolios, and students self-
assessment” (1996, p. 4).

Classroom based assessment is another term, similar to authentic assessment, that
has become more widely used. Classroom based assessment can be considered the types
of assessment that are developed by the teacher with the context of the curriculum and
requires a systematic approach to inform instruction and document student learning”
Most classroom based assessment includes activities that are authentic, such as anecdotal notes, checklists and student conferencing. It can also include tests and quizzes that teachers develop.

"Formal assessment measures include standardized tests, such as group-administered standardized achievement tests in reading, language arts, and mathematics" (Peregoy & Boyle, 1998, p. 105). Their administration, format, content, language, and scoring procedures are the same for all participants (Wilde, 2004, p. 3). Most people think of norm-referenced tests when considering standardized tests. These tests are developed by administering the test to students in a large area and calculating averages among the participants. Student scores are then compared against the "norm group" involved in the test development. Most commercially developed tests that generate student results by grade equivalents, stanines or percentiles are norm-referenced.

Criterion-referenced tests are different from norm-referenced tests, in that they are based on a specific criterion, as opposed to the scores of group of other students. Student achievement is measured by the criteria against which the test is aligned. The tests that states have developed to assess student progress against state standards are criterion-referenced tests. Students are not compared against each other or national norms, but rather the standards set by each state. Student results are usually reported against a rubric based on the standards.

Several terms related to assessment that have become more common in recent years include: "large scale assessment", "formative" and "summative" assessment. Large scale assessments, as the name infers, are assessments that are conducted with a large number of students at the same time. These assessments are usually tests required
by a state or school district to meet federal or programmatic requirements. They are administered in a consistent format, require a level of security and are usually scored as a group. Large scale assessments are normally considered summative assessments, which show outcomes and achievement, and are used for evaluation and accountability. Formative assessment is ongoing assessment primarily used to inform instruction and adapt curriculum.

Accountability

A key aspect in a discussion of assessment these days is the concept of "accountability". Accountability is discussed in conjunction with testing, assessment and student achievement. Linse, Vialpando, and Yedlin explain accountability in reference to NCLB requirements as "the process of holding individuals and institutions responsible for the strengths or weaknesses detected through assessment" (NCLR, 2005, p. iv). Linse, Vialpando and Yedlin describe the elements of a successful accountability system.

To construct a productive accountability system, assessments that are appropriate for all students need to be selected, and user-friendly and accurate reporting structures for collection and distribution of data must be developed. The roles and responsibilities of individuals charged with making decisions on instruction and resources should also be clearly delineated. (2005, p. iv)

English Language Proficiency

As discussed, the concept of English language proficiency is a complex issue that varies according to theory and point of view. Exploring this construct from the teachers' point of view is part of the focus of my research. The review of literature in the field in Chapter II will include more information on this topic.
All definitions of English language proficiency include descriptions of successful use of English in different contexts. Peregoy and Boyle provide a definition that can serve as a general understanding of the concept.

In general, language proficiency may be defined as the ability to use a language effectively and appropriately throughout the range of social, personal, school and work situations required for daily living in a given society. In literate societies, language proficiency includes both oral and written language. For our purposes as educators, we want our students to become competent in four language processes: listening, speaking, reading, and writing. (1998, p. 34)

The construct of English language proficiency will be further defined in the next chapter, explored in my research and addressed in the final interpretation of my findings.

Terms Related to Federal Legislation

Several phrases related to assessment found in NCLB include the concept of “adequate yearly progress” (AYP) and “annual measurable achievement objectives” (AMAOs). Both these phrases refer to the progress students must make according to formulas set by state education agencies on the state large scale tests. AYP is the phrase used for the progress made on the academic achievement content tests and AMAO is the phrase used for the progress made on the state English language proficiency test.

The concept of AYP will frequently surface in the data and discussion of findings. AMAO will not surface to the same extent. It is included here because it is the term used in the federal English language proficiency accountability system and the topic of the research study is English language proficiency and should be referenced as part of background information.
Summary

The assessment of ELLs in our nation is an increasingly complex and significant issue for a number of reasons. States and school districts are under more pressure to demonstrate success in meeting educational goals for all students. This pressure becomes more significant with students from multicultural and non-English backgrounds since these students have historically been underrepresented in assessment and accountability. Large scale, state controlled assessments are mandated by federal legislation as part of accountability systems.

ELLs are included in a separate, but similar accountability system developed to determine progress in English language proficiency. English language proficiency assessment as a professional field is still relatively new, lacking a strong research base to guide policy makers and educational decisions.

This research project is designed to look at the issues surrounding assessment of ELLs from the teacher's perspective. It explores what teachers are currently doing to assess their students, what they believe the assessments reveal about students and what the relationship is between large scale assessment and teacher assessment. The issues are explored from a theoretical foundation that exemplifies the complex nature of the constructs of language proficiency assessment.

Organization of Study

The study is organized into five chapters, appendixes and a reference section. Chapter I provided an introduction to the study, which included the research problem and questions. Chapter II presents a review of literature in the area of assessment and English language learners. Chapter III addresses the procedures, including the research design,
participants and other information related to the research. The findings are presented in Chapter IV. Chapter V provides a discussion of the findings, with interpretations and assertions. The appendixes, which provide copies of consent forms used, research tools and protocols, and examples of questionnaire data, can be found after Chapter V. The list of references included in the study are at the end of the document.
CHAPTER II

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Introduction

The subject of my research study, involving assessment and accountability for ELLs, is both narrow and broad in scope. The field of ELL instruction and assessment is relatively new and my specific focus of research involving the perspective of teachers is unique, limiting the availability of literature on the subject. However, many areas in education provide information on the topic. Therefore, I reviewed a broad area of education in order to define my research statement, problem and questions. Glesne advises “in conducting your literature search, cast a wide net. Do not confine yourself to your topic, nor to your discipline” (1999, p. 20).

The literature I reviewed assisted me in developing the research questions which include the following:

1. What assessments are English as a Second Language (ESL), English Language Learner (ELL), and bilingual teachers using to document the development of English language proficiency of English Language Learners?
2. What do the teachers say the assessments reveal about the students?
3. What is the relationship between teacher developed assessment information and large-scale assessment data and accountability?
Most important, though, is that the literature review assists in establishing a context for the study and research questions. Maxwell (1996, p. 25) references Miles and Huberman in his discussion of building a conceptual framework. Miles and Huberman explain that the conceptual framework explains "either graphically or in narrative form, the main things to be studied" (1994, p. 18). Maxwell cautions qualitative researchers not to simply summarize "some body of empirical or theoretical publications" (1996, p. 26).

Simply summarizing various publications is not appropriate since it can lead to a narrow focus of the literature. An approach that routinely summarizes various books and articles tends to generate a strategy of covering the field rather than focusing on theory. It can make the task seem as if it is simply descriptive. Maxwell (1996) emphasizes the purpose of literature in qualitative research as critical, along with descriptive, with an overall emphasis on establishing a conceptual framework and building theory.

Chapter II includes a description and review of the literature I used, and serves to provide a "tentative theory of what is happening and why" (Maxwell, 1996, p. 25). Maxwell explains that "The function of this theory is to inform the rest of your design—to help you to assess your purposes, develop and select realistic and relevant research questions and methods, and identify potential validity threats to your conclusions" (1996, p. 25).

I conducted an initial review to establish my research design and plan. I did not complete my review of literature once I had my plan in place. I continued to review articles, books, reports and other materials as I collected and analyzed my data, which is consistent with qualitative research (Glesne, 1999, p. 20). As patterns and themes
emerged from the data, I kept an eye out for literature that both supported and
contradicted my findings. As I began to make some tentative interpretations, I looked for
articles and books that related to those findings and interpretations. I also reviewed those
that conflicted with my theoretical basis. Glesne (1999) suggests reading conflicting
literature, along with literature that supports the research topic.

Eisenhardt (2002) also stresses the importance examining a variety of literature,
including literature with conflicting findings. The use of the conflicting literature
ultimately increases the confidence in the research findings because it pushes the
researcher to question and think more deeply about the subject. “The result can be deeper
insight into both the emergent theory and the conflicting literature, as well as sharpening
of the limits to generalizability of the focal research” (Eisenhardt, 2002, p. 24).

Reviewing literature after I collected my data helped in the analysis and
interpretation of findings. Eisenhardt explains that theory building involves “comparison
of the emerging concepts, theory, or hypotheses with the extant literature” (2002, p. 24).
She explains that the researcher looks at “what this is similar to, what does it contradict,

I used a variety of sources to locate literature. Search engines and electronic
databases, such as the Education Resource Information Center (ERIC) were most helpful.
I also subscribe to both traditional journals and newsletters, along with electronic
newsletters, and other communication systems that cite research articles and new
publications in education. While the review I conducted by no means could be
considered a comprehensive list of all publications, reports and articles in the areas of
assessment, ELLs and teacher practices, it is a selective description of significant works that provide background knowledge and information for the study.

In summary, the purpose of Chapter II is to:

- Document the importance of the research problem, (Creswell, 2002, p. 87),
- Provide information on the literature that informed my research design and data collection tools and,
- Provide a basis for the theoretical framework of the study.

The chapter is organized into trends and themes (Glesne, 1999, p. 21). First of all, I address general information on learning and language theory. Literature on mainstream assessment practices is addressed. English language proficiency assessment is then addressed. Information on current assessment practices, including legislation and policy follows, along with a review of related research on current practices. Finally, research related to large scale assessment, teacher impact, minorities and related issues is also examined.

Initially, my exploration of literature did not extensively involve research on issues related to teacher perceptions, viewpoints and attitudes. There are several reasons for this initial lack of focus in this area. A minimal amount of searching allowed me to conclude that this area represented a gap in the field. Also, the texts on conducting qualitative research are cautionary on extensive review of literature prior to the collection of data (Benz & Newman, 1998; Glesne, 1999), because of the importance of entering the field without preconceived notions.

After I collected and analyzed my data I reviewed literature related to teacher attitudes, including studies that both supported and did not support my findings. This
approach is consistent with the literature on qualitative research that “conflicting literature represents opportunity” (Eisenhardt, 2002, p. 24). These studies will be included in this chapter, with a further review later in the discussion on my findings.

Learning and Language Theory

Integral to the discussion of assessment and English Language proficiency is a review of language learning theory, with a focus on second language acquisition. Understanding first and second language acquisition is essential to understanding how to assess and evaluate progress and determine proficiency for children at different ages and grade levels. Along with language acquisition theory, learning theory must also be included since language acquisition theory has roots in learning theory.

Learning Theory

Understanding the process of learning in children is as complex as understanding what makes us human. Research and philosophical exploration related to how children learn has been addressed by a number of people. It includes the work of individuals such as Rousseau, (trans., 1979) who observed the learning processes of one child, Emile, in the eighteenth century and described how he learned and Dewey who looked at the social aspects of learning (1944) at the beginning of the twentieth century. More recently, learning theories have been developed from the work of Skinner, who looked at learning behavior as responding to external stimulus, along with Piaget and Vygotsky who provided different perspectives as to how children learn.

Learning theories can be characterized very simply in two contrasting traditions that can be described as “Scientific Management” and “Progressivism” (Putney & Wink, 2002, p. 3). Scientific Management school of thought looked at the classroom as a
workplace, emphasizing efficiency and receptive skills for students, such as listening and responding with appropriate behavior. Learning is understood as a response to outside stimulus in this viewpoint. Outcomes-based education, standards, and assessments that yields quantifiable information, are typically viewed as programs based on these theories.

“The historical roots of the management approach eventually took on other names in schools: behaviorism, positivism, traditional and back-to-basics” (Putney & Wink, 2002, p. 3).

Progressivism was popularized at the turn of the last century, primarily because of the work of Dewey (1944), and contrasts with the scientific management school of thought in that it focuses on the child’s role in learning. Progressivism supports an approach to learning that emphasizes collaboration and discovery. Dewey considered community and social interactions as important aspects of learning for children. He restructured classrooms to allow students to interact, work cooperatively and explore learning activities together.

Constructivism has its roots in progressivism. Constructivists look at learning as highly complex and multifaceted. They see learning and language development as, first and foremost, a very human activity. Just as some human behavior can form patterns and predictable activities, child development, including language, has predictable stages.

Yet, there is much in growth and development that cannot necessarily be predicted or determined or completely understood. Thus, the need to understand the student as an individual, who learns through his or her own unique skills, abilities, and way of constructing meaning, is integral to an understanding of learning. “The most important contribution of the constructivist model to instruction is its focus on the
learner’s active participation in constructing meaning rather than passive acquisition of reading and composition skills and knowledge” (Dixon-Krauss, 1996, p. 18).

Vygotsky, a pioneer in social constructivist theory, proposed an interactive viewpoint of learning reminiscent of Dewey’s emphasis on the importance of experience and community in language development. Contradictory to a more technological and behaviorist point of view characterized by the Scientific Management theories, Vygotsky believed that the child’s reason was socially constructed through interaction with adults and peers. The development of higher cognitive functions was a mediated activity, which occurred first during social interaction and that language “carries with it the meanings and intentionality of those who came before us and who now use the same tool to make meaning with us” (Putney & Wink, 2002, p. 30).

This social constructivist view of learning and language acquisition looks at child development holistically and emphasizes the problem solving and discovery that children carry out individually, but also within a group. “Social constructivism emphasizes the importance of social interaction and cooperative learning in constructing both cognitive and emotional images of reality” (Brown, 2007, p. 12).

First Language Acquisition

Language learning theory has its roots in learning theory. Basic to the discussion of language learning is an understanding of what language is. Francis and Rivera discuss the complexity of language in stating:

Language is unique among human capabilities in the roles that it plays in symbolically representing, both internally and externally, our knowledge and experience, goals and aspirations, and feelings and emotions – and in its power to create new knowledge and experiences for ourselves and others. (2007, p. 12)
Brown discusses features of language, consolidating concepts from a number of definitions, and finding eight key characteristics. These include:

1. Language is systematic.
2. Language is a set of arbitrary symbols.
3. Those symbols are primarily vocal, but may also be visual.
4. The symbols have conventionalized meanings to which they refer.
5. Language is used for communication.
6. Language operates in a speech community or culture.
7. Language is essentially human, although possibly not limited to humans.
8. Language is acquired by all people in much the same way; language and language learning both have universal characteristics. (2007, p. 6)

Brown (2007) explains that though the eight statements provide a concise definition of language, “the simplicity of the eightfold definition should not be allowed to mask the sophistication of linguistic research underlying each concept. Enormous fields and subfields and year-long university courses, are suggested in each of the eight categories” (2007, p. 6).

Lightbown and Spada (2007) discuss the fact that first language acquisition has a “high degree of similarity in the early language of children all over the world. Researchers have described developmental sequences for many aspects of first language acquisition” (p. 1). Because of the extensive research in the area of first language acquisition and development, there is considerable knowledge of what the stages of language development look like. More controversial they explain, “are questions about how this remarkable development takes place” (2007, p. 10). They describe three main theoretical positions, which include behaviorist, innatist and interactional/developmental (2007, p. 10).
Behaviorism as applied to language learning emphasizes the fact that children imitate and are reinforced positively for correct communication. Language learning is seen as developing habits of correct language usage. Behaviorists understand learning as responses to external stimuli.

The innatist perspective on language acquisition is influenced by the work of Chomsky and supports a view of language development as fundamentally innate and explains why children from different language backgrounds go through similar stages of language acquisition. Chomsky argued that humans are hard wired with a “language acquisition device” (LAD) and there is a sort of universal grammar that prevents children from pursuing wrong hypothesis in problem solving in language development.

Chomsky’s view of language and the mind reverses priorities. For him, human languages are not expressions of culture and society – in effect, human artifacts. They are, in a sense, expressions of our genes: all the existing and possible natural languages (not technical symbol systems, such as those found in the sciences) are biologically encompassed within what he calls “Universal Grammar.” If there is any dependency between language so conceived and society, culture, etc., it cannot make culture the condition of language. If anything, culture (etc.) depends on language. (McGilvray, 2005, pp. 1)

The interactionalist/developmental perspective of language development incorporates some aspects of behaviorism and innatism, but goes beyond them to emphasize the ability of children to interrelate with others and their environment. Piaget and Vygotsky are associated with this school of thought. Interactionalist/developmental theories of language would also be associated with constructivism and social constructivism. Brown explains that “In many ways, constructivist perspectives are a natural successor to cognitively based studies of universal grammar, information processing, memory, artificial intelligence, and interlanguage systemacity” (2007, p. 14).
Brown defines two branches of constructivism: cognitive and social. “In the cognitive version of constructivism, emphasis is placed on the importance of learners constructing their own representation of reality” (2007, p. 12). As discussed, he explains that social constructivism places more emphasis on social interaction and cooperation (2007, p. 12).

Second Language Acquisition and Proficiency

Second language acquisition research and theory builds on first language acquisition research and theory. Patterns of second language acquisition follow those of first language acquisition. Learners are found to go through predictable stages, supportive of Chomsky’s LAD and ideas of universal grammar. There is evidence that children and adults learn second languages through imitation and reinforcement as behaviorism suggests. Second language learners also problem-solve and interact with their environment. Similar to small children learning a first language, second language learners have the ability to creatively construct new meanings and apply language structures in situations that had not been used before.

Second language acquisition is also influenced by many factors, including age, motivation, personality, cognitive ability, learning styles, cultural background and previous with new language learning. Some of these factors, such as personality and cognitive ability, can be found in first language acquisition. Other factors, such as age, motivation and previous experience learning language are unique to those learning additional languages.

Brown explains that constructivist theory supports the multiple factors that are involved in second language acquisition. Constructivist perspectives, he explains involve
a variety of understandings, including grammar, processing of information, use of memory and intelligence along with the social activities related with dialogue and conversation (2007, p. 15). Vygotsky’s theories of language development and learning are particularly applicable for second language acquisition because of his emphasis on social interaction and the zone of proximal development (ZPD). ZPD can be defined as the distance between the actual developmental level of a child as determined by independent problem solving and the level of potential development as determined through problem solving under a teacher, another adult or peers (Chaiklin, 2004; Vygotsky, 1986).

Second language acquisition continues to be a controversial area to research because so much is yet unknown. As Lightbown and Spada explain:

Researchers and educators who are hoping for language acquisition theories that give them insight into language teaching practice are often frustrated by the lack of agreement among the ‘experts’. The complexities of second language acquisition, like those of first language acquisition, represent puzzles that scientists will continue to work on for a long time. Research that has theory development as its goal has important long-term significance for language teaching and learning, but agreement on a ‘complete’ theory of language acquisition is probably, at best, a long way off. (2007, pp. 49-50)

English Language Proficiency

Intrinsic to an understanding of second language acquisition and development is an understanding of what it means to be proficient in a language. At a certain point, a student who is learning English must be considered to have mastered sufficient areas of the language for his age or grade level so that he or she no longer is considered “limited” in English. Just as understanding the subject of how languages are learned is complex, the issue of language proficiency is also complicated.
The issue of proficiency in English becomes more difficult when considering children, since an aspect of childhood is the development of language skills. Children are in the process of learning all things, including language. Children learning English as a second or other language are compared with their English speaking peers who are also developing oral and literacy skills in the language. The “bar” for proficiency continues to move. Another issue that complicates our understanding of proficiency in the English language is our changing world. The needs of the work force have changed due to advances in technology.

Our understanding of competency for all English users has changed from the past. This change has been greater in the area of reading and writing. The understanding of what it means to be literate in the 21st century has changed as the modern technological, communication-driven workforce requires complex literacy skills as compared to the agrarian based economy of the past.

Gordon and Gordon (2003) explain that the concept of literacy as a skill necessary for the ordinary individual grew as Europeans settled on the American continent and brought their literacy heritage with them. “With each passing European generation, literacy became a more important means for the upward social mobility of individuals and families” (p. 3). In the early days of the United States the level of literacy for an economy that was based on farming was minimal by our current expectations.

Neither did the Native Americans at the time of European contact have a strong need for literacy. Their tradition was oral. When Trout explains that “Native American literature is the oldest linguistic legacy in North America” (1999, p. xvii), she means oral literature. She explains the importance of that oral tradition.
In a world where myth and mystery united, rituals linked the spirits of hunters and animals. Sacred stories formed the heart of ceremonies. In a precarious life cycle, fathers and mothers pictured family memories and set cultural lessons in stories for their children. (1999, p. xviii)

Trout explains that when the European conquest and colonization put Native American lives and languages at risk, many decided to become proficient in English. Literacy became a greater priority than it had in the past. "Some natives learned that the written word was powerful in dealing with literate Euro-Americans" (1999, p. xix) she explains. Along with the forced English language instruction of the boarding schools, the desire to protest injustice was a reason Native Americans developed levels of literacy that allowed them to communicate effectively. Personal expression was also a motivation for literacy. "As Indians wrote and told about their experiences on reservations and off, they created autobiographies within diverse tribal contexts" (Trout, 1999, pp. xx-xxi).

Other ethnic groups have gone through similar experiences with English as the Native Americans. The ability to read and write well in English provides people with the skills to compete in the job market and other areas. Gordon and Gordon provide a working definition of literacy for today's world as "the degree of interaction with written text that enables a person to be a functioning, contributing member of the society in which that person lives and works" (2003, p. xv).

As the need for higher levels of literacy grew for all people, the understanding of proficiency in English for second language learners evolved. Research demonstrated that basic, context reduced language was learned more easily than abstract language. Cummins (1984) first made the distinction between basic interpersonal communicative skills (BICS) for social interaction and the cognitive academic language proficiency
(CALP) needed for success in the classroom. He explains that his purpose was to make a point that "it takes language minority students considerably longer to attain grade/age-appropriate levels of English academic skills than it does in English face-to-face communicative skills" (1984, p. 152). Cummins' theories of BICS and CALP have come to be commonly accepted by educators as a way to explain children's language development.

Gottlieb adds to the discussion of language proficiency and the need for literacy skills. She explains the differences between a basic English language proficiency and proficiency that also encompasses academic language skills. "Overall, language proficiency represents general knowledge and language use in the areas of listening, speaking, reading, and writing" (2003, p. 12). She explains that academic language is:

primarily centered on language use in contexts and interactions in and outside of school. In contrast, academic achievement typically represents subject matter knowledge, skills, and concepts in the areas of language arts, mathematics, science, and social studies. It is an expression of conceptual learning that is directly linked to school-based curricula and, in recent times, academic content standards. (2003, p. 12)

Gottlieb uses a visual to express the blending of language proficiency and academic achievement to produce a system for English language proficiency. Figure 1 demonstrates how the concept of academic language proficiency includes both elements from traditional concepts of language proficiency and the language needed to succeed in academic achievement.
Researchers have continued to discuss the concept of academic English language proficiency. Butler, Castellon-Wellington, and Stevens discuss the different theories of academic language, identifying a view that "defines academic language primarily in terms of the language functions and corresponding structures that students must use in the classroom" (2000, p. 5), and a second model that "proposes a distinction between academic language and social language, with an emphasis on context and cognitive difficulty" (p. 5). Current definitions integrate the two views, according to Butler, et al. They cite a third view of academic language as proposed by Solomon and Rhodes which sees academic language as a "register that includes task specific stylistic registers" (2000, p. 6). Butler et al. discuss the variety of different activities involving different styles and registers that are implicit in academic language tasks.

Reading a textbook, discussing a poem, or taking a standardized content assessment are all examples of academic tasks that may require variation in register or the stylized registers discussed by Solomon and Rhodes [1995]. In other words, these tasks may require the use of specific discourse, functions, vocabulary, and/or structures not used across other academic tasks. (2000, p. 7)
Bailey examines academic English language proficiency in terms of the language demands placed on students in public schools. She builds on the work of earlier researchers, such as Cummins, Chamot and O’Malley and supports the concept of the need for a definition of academic language that encompasses language used for learning new information, describing abstract concepts and other skills used in the classroom (2007, p. 9). She diverges from Cummins’ understanding of BICS and CALP, though, by emphasizing the importance and complexity of social language.

Think about the complex and highly sophisticated social applications of language that are needed to woo or deceive a loved one; and conversely, there are contexts in which AEL can be as simple as responding with a head nod to yes or no question....In some regard, it is not meaningful to conceive of language as either social or academic, rather it is the situation that is either predominantly social or academic. Even this distinction is problematic. Dewey’s pedagogic creed (1897) serves as a good reminder that schools are simply one more social context or ‘community’ and that language is always ‘a social instrument.’” (2007, p. 9)

Those who believe in the importance of bilingualism and the preservation of home languages believe that the child’s first language plays an important role in second language and English acquisition. They must not be considered separately. Falchi, Garcia and Kliefgen explain this role.

It may seem counter-intuitive to support the use of a child’s first language in the process of helping that child achieve to a high level in an English-language school system. But the benefits of such practices are explained by the concept of linguistic interdependence, which means the two languages bolster each other and thus the student in his or her acquisition of knowledge. (Garcia, Kliefgen & Falchi, 2008, p. 27)

Falchi et al. also address the issue of BICS and CALP as an oversimplification of the complexity of English language proficiency. They support the theories of Cummins on the benefits of bilingualism enhancing English language proficiency, but see language
proficiency as much more multifaceted than the differentiation of language proficiency into social English and academic English. They explain that:

an alternative framework rejects a binary view of language and suggests that both BICS and CALP are over simplified terms. This alternative framework recognizes that the linguistic repertoire is a complex phenomenon comprising multiple codes and modes or channels of expression and that language is contextual. (Zamel & Spack, 1998) (2008, p. 28)

Falchi, et al. emphasize the importance of the community outside of school in the development of English language proficiency as Bailey and others do.

In other words, learning academic language is not a neutral activity, easily divided into two modes of communication – spoken and written. Rather, as recent scholarship has shown, learning academic literacy entails much more: full academic literacy requires skills that are multimodal – spoken and written modes intricately bound up with other visual, audio and spatial semiotic systems (Jewitt & Kress 2003; Kleifgen, forthcoming; Kress, 2993: New London Group, 2000). These literacy scholars note that the acquisition of these complex technical skills is contingent upon wider societal factors beyond school. (2008, p. 28)

Cummins, though, never intended it to represent to view language development only in this simplified framework. He explains that “dichotomizing ‘language proficiency’ into two categories oversimplifies the phenomenon and makes it difficult to discuss the developmental relationships between language proficiency and academic achievement” (1984, p. 152).

Francis and Rivera summarize many of the definitions in stating that “Language proficiency involves the effective use of language to accomplish different objectives of importance to the language use, and reflects linguistic competencies in multiple dimensions” (2007, p. 18).

As in all areas of education, research can often conflict with practice. At the same time researchers are struggling with a valid definition of English language proficiency
that encompasses language learning theory and the needs of individuals in the world today, educators are making decisions on a daily basis on whether students have attained proficiency or not. There are many different standards in states, classrooms and among teachers for considering when a student is proficient. Linquanti (2001) discusses how reclassifying students to proficiency can be based on policy, politics, and funding, rather than concrete evidence of performance. Educators may look first at the fact that there may be some benefit to the school district to moving a student out of the category of ELL. He addresses the ramifications of exiting students too soon from the category of ELL or "Limited English Proficient" to "Fully English Proficient". Decisions may not be made completely on student ability.

Part of the tension surrounding reclassification has to do with what educators, policymakers, and the public understand to be the benefits, risks, and meaning of students being classified in particular language categories. How do LEP students profit – or not – as a result of their classification? What do FEP students gain – possibly lose – as a result of their reclassification? (2001, p. 7)

Linquanti explains that though students benefit from services by the LEP classification, there is also the risk of stigmatizing students by maintaining them as LEP for a long time. "On the other hand" he states, "reclassifying English Learners prematurely when they lack needed academic language skills or content-area knowledge and abilities also puts them at risk" (2001, p. 7).

The definition of English language proficiency continues to be debated nationally as the issue becomes more controversial and involves educational decisions that have impact on children and schools. Defining proficiency and the related concepts of progress towards proficiency, and "exiting" students, are key pieces of assessment and accountability systems. Wolf, et al. found "substantial variation" (2008c, p. 3) among
state and school districts policies on definitions of proficiency and comparability of
assessment results in a recent review of state practices.

On the whole, English language proficiency from a social constructivist point of
view involves a complex picture of an individual who not only has acquired a set of skills
and strategies in oral and written language that allows him or her to access knowledge to
be successful in the classroom. A definition of language proficiency would involve
creatively using problem-solving strategies for language in different circumstances. It
would include the use of language in a group or community situation. It would also
involve the integration of the individual’s background and culture and role in his or her
current world.

Educational Assessment

Educational assessment is a broad field, encompassing a variety of theories,
concepts and practices. It is influenced by policy and research. It is driven by a need to
find out how students, programs and institutions are doing. Ultimately it is influenced by
the educators that implement it and use it.

Background

Just as our understanding of literacy and English language proficiency has
changed and evolved, our understanding of assessment has changed and evolved. As
literacy demands in society have increased and changed, assessment has changed.
Assessment practices have grown from classroom quizzes and dictations administered by
schoolteachers to large-scale standardized tests.

Dewey advocated a child-centered approach to assessment. Dewey continued the
tradition of Rousseau, the 18th century French philosophy who emphasized the
importance of allowing a child to develop and explore the world naturally. Dewey states that: "education is a process of development in accordance with nature" (1916, p. 112). He cautions against taking an industrial approach to education and assessment. "The present industrial constitution of society is, like every society which has ever existed, full of inequities" (1916, p. 119). Dewey explains that progressive education, as opposed to industrialist approach supports equity for all students.

It is the aim of progressive education to take part in correcting unfair privilege and unfair deprivation, not to perpetuate them. Wherever social control means subordination of individual activities to class authority, there is a danger that industrial education will be dominated by acceptance of the status quo. Differences of economic opportunity then dictate what the future callings of the individuals are to be. (1916, p. 119-120)

Similar to learning theory and educational philosophies, assessment policies and practices reflect a social progressive view of instruction or a more technological, behavioral approach to instruction and assessment. Standardized, norm-referenced testing primarily developed from a behaviorist model. Policy, politics and reformation movements also influence assessment. Stoynoff and Chapelle explain:

The educational landscape in the United States changed dramatically in the early 1980s. The U. S. educational reform movement was precipitated by the publication of *A Nation at Risk: The Imperative for Educational Report* (National Commission on Excellence in Education, 1983). This 32-page report to the U. S. secretary of education documented a decline in the academic quality of U. S. educational institutions (public and private, from kindergarten through university), and it recommended five major reforms to correct the declines in achievement. (2005, p. 4)

Stoynoff and Chapelle explain that two of the recommendations were “to restore an academic core (called new basics) to the curriculum” (p. 4) and “to implement more rigorous and measurable standards for academic performance” (p. 4). These goals
involving increased emphasis on academics and measurable standards were incorporated into educational legislation.

Currently, an educational philosophy more consistent with the scientific management, behaviorist or technological philosophies and theories dominates. The emphasis on uniform standards, benchmarks, progress goals and outcomes is representative of these approaches.

*General Education Assessment*

Defining assessment continues to be influenced by educational theory and philosophy. The International Reading Association (IRA) and the National Council of Teachers of English (NCTE) address the changing history of assessment by discussing the transmission view of knowledge, curriculum, and assessment. “In a transmission view, it made sense to develop educational standards that specified the content of instruction before developing assessment procedures and engagements” (1994, p. 5). The transmission view can be seen as consistent with the scientific management, behaviorist approaches, as opposed to an inquiry approach, consistent with progressivism and social constructivism. IRA and NCTE explain that the shift from a knowledge transmission philosophy to inquiry changes the role of assessment.

Quality assessment, then, hinges on the process of setting up conditions so that the classroom, the school, and the community become centers of inquiry where students, teachers, and other members of the school community investigate their own learning, both individually and collaboratively. The onus of assessment does not fall disproportionately upon students and teachers (which is often the case in schools in our society today); instead all those involved in curriculum inquiry are held responsible for investigating the roles they have played. (1994, p. 6)
Hill, Ruptic, and Norwick who primarily address alternative and classroom assessment describe assessment as "the process of gathering evidence in order to document learning and growth" (1998, p. 15).

Within the concept of assessment is a broad range of activities. The assessment vocabulary these days includes large-scale, classroom based, authentic, formative, summative, commercially developed, locally developed, norm-reference or criterion-referenced, and a variety of other terms. These terms describe the various approaches to instruments and activities that provide information on student learning and achievement. In many cases definitions of assessments overlap.

Assessment terminology can vary and traditional definitions have become blurred. For example, standardized assessment is a commonly used term, but can actually mean different things. Standardized assessment has traditionally meant formal, commercially-developed, norm-referenced tests. Standardized means that the "administration, format, content, language, and scoring procedures are the same for all participants – these features have been 'standardized'" Wilde, 2004a). Most large-scale assessments can be considered standardized, but they are not necessarily norm-referenced, or commercially developed. Most of them are "standards-based" which fits into the category of criterion-based assessment. On the other hand, many classroom or locally developed assessments can include standard methods of administrations and scoring, allowing them to be included in the definition of a standardized test. Portfolios, for example, can follow a standard method of format.

Formative and summative are two terms that have become more commonly used to differentiate between different types of assessment. Formative assessment involves
assessment that is more process oriented, focusing on assessing growth and development, while summative assessment "aims to measure, or summarize, what a student has grasped" and implies "looking back and taking stock of how well that student has accomplished objectives, but does not necessarily point the way to future progress" (Brown, 2004, p. 6). Formative and summative assessment can be formal or informal.

Another area of distinction in assessment are the terms that differentiate between assessment that fits along a paradigm of informal, alternative, "home-made" or teacher developed and a paradigm that is more formal, standardized and structured. As in the other areas, the terms are not always appropriate and can overlap, but for lack of better terminology, the words, informal, and formal, will be used.

Informal assessment can include a broad range of activities, including observations, skills, checklists, portfolios, conferencing, peer reviews, and self-assessment. Authentic assessment would fit into the category of informal assessment. As discussed in the definitions in Chapter I, the purpose of authentic assessment is to connect assessment to real situations as close as possible. Because of this close connection between the task and the assessment situation, authentic assessment overlaps with performance assessment. Likewise, many classroom assessments are considered authentic and performance based. For example, a teacher may incorporate many ongoing literacy assessments, into her curriculum such as writing and reading conferences, observations, checklist and rubrics for student work. These classroom assessments are based on authentic activities.

Authentic, performance and classroom based assessments are usually used in a formative manner, to inform instruction and provide ongoing information on student
progress. They can also be used once a year or at the end of an instructional period as a summative measure also.

Formal assessment is usually seen as standardized or norm-referenced testing (Cabello, Fisher, Flood, & Lapp, 2001, p. 10). Criterion-referenced tests are included in lists of traditional, formal assessments and informal, alternative assessment. There are a number of assessment areas where there is not a clear distinction between traditional and alternate assessment. Computer-based assessment can involve aspects of traditional testing and performance based assessments. Likewise, the current standards-based, criterion referenced tests often include traditional activities, such as multiple choice along with constructed responses, involving student writing, which is considered more of a performance assessment. Most formal assessment measures are used in a summative manner. They can provide information on how a system, group of students or program is doing.

As addressed in Chapter I, informal and formal will be used as categories to include differentiate between those assessments that follow a more standardized approach to administration. Informal assessment will include assessment that allows more flexibility in administration.

Assessment varies according to needs and purposes. It can be large-scale or individual. It can be considered high-stakes if the results have significant ramifications for an individual or group. Typically, large-scale and high-stakes assessment includes some form of testing.

Many in education have come to associate the term assessment with testing. Technically, assessment is a broader term. It is not just testing. “Tests,” as Brown
explains, “are a subset of assessment” (2004, p. 4). Gottlieb explains that “testing, by being one data source, contributes to assessment” (2003, p. 3). Linn and Gronlund explain the differences between testing and assessment.

Assessment is a general term that includes the full range of procedures used to gain information about student learning (observations, ratings of performances or projects, paper-an-pencil tests) and the formation of value judgments concerning learning progress. A test is a particular type of assessment that typically consists of a set of questions administered during a fixed period of time under reasonably comparable conditions for all students. We sometimes speak of testing and assessment together even though tests are a specific type of assessment. (2000, p. 31)

Most states refer to their large scale standards-based tests as assessments. For example, the instrument North Dakota has developed to meet requirements of Title I of NCLB is called the North Dakota State Assessment (NDSA). In fact, NDSA is a paper and pencil test. While this distinction may seem to be a minor issue, it becomes significant when looking at the whole picture and the impact on children in schools in the state. At test is a single activity, providing one data point. The term, assessment, infers a broader range of data. Using the term, assessment, for a high stakes test, gives the impression that decisions are made, based on a wider criteria than one data point.

Testing is a broad concept encompassing many types of measurements or activities. Brown discusses the different types of tests used in the instruction of culturally diverse students, including language aptitude tests, proficiency tests, placement tests, diagnostic tests, and achievement tests (2004, pp. 42-48). In fact, this is only a small number of the types of tests children will confront in their educational careers. There are intelligence tests, readiness tests, reading tests, and various other commercial, standardized tests used in schools, in addition to teacher-prepared, classroom tests.
Brown defines a test as “a method of measuring a person’s ability, knowledge, or performance in a given domain” (2004, p. 3). Many tests used in education are commercially developed by testing companies and have standard methods of administration.

As discussed in Chapter I, tests are typically divided into norm-referenced tests and criterion-referenced tests. “When referring to standardized assessments, most people think of norm-referenced tests (NRTs). NRTs typically are used to sort people into groups based on their assumed skills in a particular area (for instance, those in the top 10% of skills)” (Wilde, p. 2004a, p. 3).

Brown explains that in “norm-referenced tests, each test-taker’s score is interpreted in relation to a mean (average score), median (middle score), standard deviation (extent of variance in scores), and/or percentile rank” (2004, p. 7). He explains that the purpose of norm-referenced tests “is to place test-takers along a mathematical continuum in rank order” (2004, p. 7).

Wilde explains that “criterion-referenced tests (CRTs) measure how much or whether specific knowledge has been gained; that knowledge is the criterion against which the participant is measured. CRTs must be aligned closely to the curriculum (which of course must be aligned closely to the district or state content standards) in order to ensure that what is being tested is what has been taught” (Wilde, 2004a, p. 4).

The role of testing and its place in education these days is controversial. Scherer explains that large scale, standardized tests “sort, track and stratify individuals and groups, separating the qualified from those judged less qualified” (2005, p. 9). She states
that “Parents, policy-makers, and teachers all look to tests as the definitive proof that students are learning” (2005, p. 9).

The type of testing that Scherer discusses, in which decisions are made about students, teachers, and educational programs is often referred to as “high-stakes” tests. High-stakes testing is a concept that is becoming more common in current educational vocabulary. In their position statement on high-stakes assessments in reading, IRA defines high stakes tests:

High-stakes testing means that the consequences for good (high) or poor (low) performance on a test are substantial. In other words, some very important decisions, such as promotion or retention, entrance into an educational institution, teacher salary, or a school district’s autonomy depend on a single test score. (2005, p. 9)

Other uses for high-stakes tests can include exiting students from services, tracking, funding and the decisions made related to the federal requirements for making adequate progress and meeting achievement objectives. Heubert and Hauser explain that the use of high stakes testing is growing. “In particular, states and school districts are using such tests in making high-stakes decisions with important consequences for individual students” (1999, p. 1). They explain that “These policies enjoy widespread public support and are increasingly seen as a means of raising academic standards, holding educators and students accountable for meeting those standards, and boosting public confidence in schools” (1999, p. 1).

The concept of high stakes testing and its emphasis on the consequences of an assessment is a significant concept. The use and purpose of assessment as a reflection of theory and philosophy is as important as the type of assessment. Whereas assessments must be based on theory, their use, purpose, and practice is also based on a theoretical
foundation. Heubert and Hauser, in their roles on the national committee for appropriate

test, use developed principles for the appropriate use of tests for educational decisions.

These principles address several issues, including:

- Tests are only valid for specific purposes and should not be used for other
  purposes.

- Tests are not perfect and no single test score can be considered a definitive
  measure of student knowledge.

- An educational decision that will have major impact on a test taker should not
  be made solely or automatically on the basis of a single test score.

- Neither a test score nor any other kind of information can justify a bad

Assessment and Theory

Many in the field of education are concerned about the increasingly important role

of testing and assessment and the incompatibility of current assessment systems with the

learning theory and educational philosophy that they support. Assessment, as discussed,

is a broad issue and includes not only the assessment instruments themselves, but also the

manner in which the assessment is administered and the use of the information. Johnston

and Costello discuss some of the issues related to assessment and policies, emphasizing

the role of assessment procedures. “‘What gets assessed is what gets taught’ is a common

assertion whose meaning is often underestimated. It is not just what gets assessed, but

how it is assessed that has implications for what is learned.” (Costello, P. &

The theoretical basis of the educational program influences how assessment is used. Assessment conducted in a setting that supports social constructivist language acquisition theories would include holistic, child-centered activities and activities that reflect the subtle varieties of language use in different domains of society, culture, and modes of learning. The assessment system would include not only skills that students have mastered, but also skills, language knowledge, and competencies that they are in the process of developing. It would involve assessing emergent and potential language, along with mastered language.

Tierney summarizes an understanding of assessment in this context. His principles for assessment come from theory and research which is a mix of “child-centered views of teaching, pluralistic and developmental views of children, constructivist views of knowing, and critical theoretical views of empowerment” (2005, p. 23). These principles, which are also used in my research as a data collection instruments, are available in Appendix B.

Tierney believes assessments should emerge from the classroom, assessment practices should be client centered and reciprocal and assessment should extend beyond improving our tests to the purposes of assessment and how results from assessment are used, reported, contextualized and perceived. He states that diversity should be embraced, not slighted. He also believes that assessment procedures need to be non-standardized to be fair to the individual, and learning possibilities should be negotiated with the students and stakeholders rather than imposed via standards and assessment that are preset, prescribed or mandated (Tierney, 2005, pp. 24-36).
The assessment of English language proficiency has followed general trends in assessment and education as a whole. The movement towards basic skills in education and rigorous standards and assessment that emerged from the last decade also influenced English language proficiency assessment. An increased understanding of second language acquisition and language proficiency has allowed for the development of better assessment tools. The field has also been influenced by the federal and state legislation and the growing numbers of ELLs. The impact of the changing policies on English language proficiency assessment will be addressed in the section on policy and legislation.

Both formal, standardized tests and informal, alternative assessments are available for the purposes of assessment of children’s language development in English. There are many different types of standardized tests used to assess English language proficiency of students from preschool to university level (Stoynoff & Chapelle, 2005). Whether they accurately assess English language skills necessary for the complex world of today is an issue of debate.

Essentially, the assessment of English language proficiency can be seen as the gathering of information of a students’ competency or level of proficiency in the English language. It usually involves all modalities of English. Depending on the theoretical basis of the assessment, it would involve English usage in different registers and domains. There are different types of language proficiency assessment, depending on the purpose of assessment (Gustke & Navarrete, 1996).

Some of the purposes of English language proficiency assessment include:

- Identification and placement of students;
• Diagnosing individual student needs (e.g., assessing developmental status, monitoring and communicating student progress, certifying competency, determining needs);
• Reclassifying within or transition from support services;
• Informing instruction (e.g., evaluating instruction, modifying instructional strategies; identifying instructional needs);
• Evaluating programs; and
• Providing accountability information. (Gottlieb, 2006; Cabello et al., 2001)

Just as language proficiency is not an easy concept to define; assessing language proficiency is not an easy task. Wilde explains that:

Assessing language proficiency is a difficult issue. Language proficiency assessments:
• Must be appropriate for students of different cultural, ethnic, social, and educational backgrounds;
• Are assumed to be able to predict how well a student will do in academic classes although they do not include information about cognitive abilities or academic achievement; and
• Tend to measure specific aspects of language (e.g., word choice, grammar) rather than overall communicative competence, which has a repertoire of communication skills that can be used in a variety of situations. (2004a, p. 5)

The difficulty of assessing English language proficiency is related to the evolution of the concept of English language proficiency. As there became greater awareness of the need for students to develop academic language proficiency, assessment measures needed to accommodate this new understanding. English language proficiency tests needed to be valid and reliable measures of academic English language proficiency. Wolf et al. explain how modern validation theory places emphasis on the definition of English language proficiency:
The validation process begins with consideration of the construct to be measured (e.g., ELP), the interpretations that are to be drawn from the test (e.g., what level of language proficiency students possess), and the purposes of a test (e.g., for placement, for determining progress, for making redesignation decisions). This basic specification is then the basis for asserting claims and directing the specific types of validity evidence that should be collected. (2008a, p. 11)

Using a test for multiple purposes that ELP tests are used for can make a difference in validity. Whereas a test may be valid for the purposes of identifying a student as ELL, it may not be valid for monitoring progress or exiting students. “Each of these purposes may require unique evidence, and a test of ELP that has been validated as serving one purpose cannot be assumed to serve another” (Wolf et al., 2008a, p. 11).

Validity and reliability are important for English language proficiency tests for many reasons. As these tests are used for high stakes decisions that impact students, teachers and schools, there is a greater need for quality. Wolf et al. emphasize the importance of quality in the tests because “if a state assessment does not accurately reveal individual students’ level of English proficiency, they may be laced in inappropriate academic environments and/or inappropriately transitioned to FEP status, which in turn may impede their subsequent progress” (Wolf et al., 2008a, p. 20).

Wolf et al. reviewed definitions of academic English language proficiency, along with the operationalization of academic English language proficiency in an assessment, and found challenges and variance (2008a; 2008b). They also found the use of an English language proficiency test as the sole criteria for exiting students to be problematic.

Ultimately, English language proficiency assessment must provide information on student’s capability with the English language. It must be based on a theory and
philosophy of language acquisition and development. It must include a definition of proficiency and provide information on where a student stands in relation to that definition. It must give the teacher, administrator, parent, or student a picture of where the student is at in terms of the definitive goals of full English language proficiency.

Impact of Culture on English Language Proficiency Assessment

From a social linguistic and constructivist approach, culture is a significant factor in understanding language. Language is embedded in a culture and culture influences language. "Language is deeply implicated with culture, and an important part of it" (Nieto, 1999, p. 60). Culture as a concept and influencing factor in education has grown and changed as our understanding of literacy and language proficiency has grown and changed. Culture has been defined in many ways. The most useful definition is a broad one that encompasses a variety of aspects. Lindsey, Lindsey, Robins, and Terrell state that:

For us, culture is a set of common beliefs and practices that a person shares with a group. These beliefs and practices identify that person as part of the group, and they help other group members to recognize that person as one of them. Most individuals identify with one or two groups very strongly – this is their dominant culture. They may also identify in a lesser way with other cultural groups. (2002, p. 52)

Lindsey et al. explain that ethnic cultures “are groups of people who are united by ancestry, language, physiology, and history, as well as by their beliefs and practices” (2002, p. 52). Culture is not static; nor are there always clear distinctions as to a child’s cultural background. Many children come from mixed cultural backgrounds in today’s mobile society.
Home or heritage language, which is closely connected with culture, is not always clearly defined either. Children of new immigrant backgrounds quickly began to assimilate English slang and everyday phrases into their vocabulary. Native American students can qualify as ELL, according to the federal definition, though English may be their dominant language and knowledge of their heritage or tribal language is minimal.

There are strong cultural connections and vast difference between Native American and European languages. For this reason, along with other socio-affective factors, a dialectical form of English has evolved on some reservations that has remnants of the syntax and structure of the tribal language. Reyhner, discussing Leap's research in a review of his book on American Indian English, explains that Leap's thesis is that ancestral tribal languages influence both the grammar and the use of Indian English dialects. Thus, each tribe has its own unique version of Indian English (Reyner, 1994). Native American students are one of the groups of students that have historically underachieved on standardized tests (Fayden, 2005; McCarty, 2005; Little & McCarty, 2006). In contrast, Demmert has found that student performance has increased in instructional programs that address culture and language (2001).

This interconnection between language and culture and the continual influence of the student's home or heritage language and culture is a key factor in the acquisition of English and understanding of proficiency. It continues to be researched and discussed as different viewpoints of culture language, bilingualism and academic achievement impact educational standards and expectations.
Historically, students from different language backgrounds were minimally addressed in educational policies and practices. Students who came from different language backgrounds were primarily expected to assimilate into the mainstream English speaking classroom with little support or accommodations. There was little understanding or attention to the process of second language development or instruction. Policies, whether overt or covert, emphasized assimilation.

"The history of access to educational language rights in minority languages, from the colonial period to the present, indicates a mixed bag of official and unofficial policies" Wiley states (2002, p. 60). He explains that the policies primarily emphasized English language use to the exclusion of the Native and immigrant languages. Bilingual education, which research has shown to be an effective method of education for non-English speaking students, was minimally used (Crawford, 1991, 1992; Crawford & Krashen, 2007; Cummins & Swain, 1984; Hakuta, 1986; Leyba, 1994). Though some immigrant groups were able to maintain their heritage language through bilingual schools "for the 19th century education establishment, linguistic assimilation was the ultimate goal for immigrant students" (Crawford, 1991, p. 21).

Early attempts by Native Americans to educate their children in their tribal language were short-lived. The Cherokee tribe had developed literacy skills in their language and developed bilingual materials to educate the young people of the tribe. These schools were successful. Crawford (1991) explains that the schools established by the Cherokee tribe in the 1850s in Oklahoma had a higher literacy levels than the white populations of Texas or Arkansas.
The program was dismantled and the formal policy of the U. S. Bureau of Indian Affairs (BIA) became a policy of only allowing English in the curriculum. This emphasis on English was an effort to repress the Native culture and assimilate the tribes. Hispanic experiences in school were similar to the Native Americans in that Spanish was primarily forbidden and students were expected to "sink or swim" in school (Crawford, 1991).

Protecting the rights of students who come from different language backgrounds was first addressed in the Civil Rights Act of 1964. The USDE Office of Civil Rights OCR described the assessment requirements regarding ELLs for districts based on Title VI of the Civil Rights Act:

Districts should:

- identify students who need assistance;
- develop a program which, in the view of experts in the field, has a reasonable chance for success;
- ensure that necessary staff, curricular materials, and facilities are in place and used properly;
- develop appropriate evaluation standards, including program exit criteria, for measuring the progress of students; and
- assess the success of the program and modify it where needed (USDE, 1998, p. 7).

The Elementary and Secondary Education Act (ESEA) was created in 1968. The legislation came not long after the Civil Rights Act and reflected interest in addressing minority issues in the nation. ESEA’s focus was to adjust the inequity that resulted from students in poverty attending minimally funded schools. Title I of the Act provided funding for schools with high levels of poverty. Title VII: the Bilingual Education Act provided funding for bilingual education programs for students from non-English language backgrounds. "For more than 30 years, the Bilingual Education Act of 1968
authorized grants to support instructional programs for ELLs. Most of the funding was set aside for approaches that used the native language and, from 1994 to 2001, for those that actively encouraged bilingualism and biliteracy" (Crawford & Krashen, 2007, p. 53).

A U. S. Supreme Court decision on a case involving a class action suit on behalf of Kinney Lau and 1,789 other Chinese students in 1974 further clarified the protections and rights language minority students have for appropriate educational services. The court ruled in Lau vs. Nichols that merely providing the same textbooks and facilities does not necessarily guarantee equal educational opportunities. School districts have a responsibility to provide appropriate educational support for students who are not proficient in the language of the curriculum.

Various court decisions and memoranda from the federal government have continued to spell out the responsibilities of school districts for language minority students. These decisions have carried the weight of legislation and allowed the U S Office of Civil Rights (OCR) authority to investigate complaints and conduct reviews of school districts.

*Instructional Programs*

School districts responded by federal requirements by developing programs for ELLs. Instructional programs, based on either English language development, bilingual education or an incorporation of both activities were used. OCR has not required a specific type of instructional program, but rather put the responsibility on the school district to demonstrate effectiveness of the model of program chosen. Bilingual education programs can be defined as:
Educational programs that use two languages, one of which must be English, for teaching purposes. Bilingual education programs have taken many forms, but two goals are common to all: (1) to teach English and (2) to provide access to the core curriculum through the home language while students are gaining English language proficiency (Lessow-Hurley, 2005). (Peregoy & Boyle, 2008, p. 23)

More common have been programs that only use English for instruction. “ESL pull-out” historically, has been the most common approach to teaching ELLs. ESL pullout can be defined as a program in which:

English learners receive the majority of their instruction in the regular classrooms alongside their monolingual English-speaking peers. However, they are pulled out of the classroom on a regular basis to receive additional help from an ESL teacher or aide. The help they receive consists of English language development activities and reinforcement of subject matter being taught in the regular classroom. The goal is to help students get by while becoming proficient in oral and written English. (Peregoy & Boyle, 2008, p. 26)

Instructional methods for ELLs have been influenced by mainstream trends and theories in education. ESL instruction has been delivered through an isolated, skills-based approach. ESL teachers have also used more holistic methods of language development. Freeman and Freeman explain the need for “whole language” for second language learners in the early 1990s when mainstream literacy practitioners were debating the use of whole language as opposed to skills based teaching methods based on behaviorist models. Freeman and Freeman state:

For those students whose first language is not English, whole language is not only good teaching, it is essential. Whole language may be the only road to success for bilingual learners. The instruction that many bilingual learners have received in schools has been for the most part fragmented and disempowering (Crawford 1989; Cummins 1989b; Flores 1982). (1992, p. 5)

Freeman and Freeman define whole language as instructional that is classroom based and grounded in research, including second language acquisition research (1992:
Theoretical foundations would include social constructivist learning theory and progressive educational philosophies.

More recently, other forms of instructional, in addition to Pull-out ESL, have become popular. There are many terms used for the various methods. Peregoy and Boyle (2008) provide three additional categories other than traditional ESL. These include “Sheltered English” or Specially Designed Academic Instruction”, “English Language Development” and “Structured English Immersion”. A common trend in these approaches is to use unique teaching strategies to assist students in accessing content material while also learning English.

Assessment Programs

As school districts and educators became aware of the unique construct of English language proficiency and ESL programs were developed, the need to assess the students surfaced. Commercial testing companies responded to the need to develop English language proficiency tests and created standardized, norm-referenced tests. Traditionally, assessment of English language proficiency focused on basic skills in English.

When Del Vecchio and Guerrero reviewed currently used English language proficiency tests for a federally funded assistance center in the mid 1990s, they found five tests that were commonly used. These tests include:

1) Basic Inventory of Natural Language (Herbert, 1979);
2) Bilingual Syntax Measure (Burt, Dulay & Hernandez-Chavez, 1975);
3) The Idea Proficiency Test (Dalton, 1978; 94);
4) The Language Assessment Scales (De Avila & Duncan, 1978; 1991); and
Del Vecchio and Guerrero explain that the purpose of the handbook they developed was not to critique the tests, but to provide information. They do provide a summary of their review and include several issues related to the tests (1995). There were fundamental differences among all the tests. Different language proficiency tests were shown to generate different levels of proficiency for the same student. Students who are classified as limited in English on one test may be classified as proficient on another test. Test developers tend to use a discrete point, rather than integrated approach to language testing. These discrete point types of testing which lends itself to paper and pencil activities, may not reflect language acquisition theory.

Del Vecchio and Guerrero explain that "Consequently, and to the degree that the above two points are accurate, currently available language proficiency tests not only yield questionable results about students' language abilities, but the results are based on the most impoverished models of language testing" (1995, p. 12).

Informal assessment, including authentic, performance, and classroom assessment has also been used to assess ELLs in the past. O'Malley and Pierce explain authentic assessment is more appropriate for ELLs because it gives the teacher a more complete picture of the student (1996). Many of the authentic assessments they describe were developed for mainstream education and adapted for ELLs. They include oral interviews, anecdotal records, story retelling, writing samples, projects, experiments, constructed response items, teacher observations and portfolios. Authentic assessment protocols often use a rubric and different from mainstream literacy assessment rubrics in the expectations.
Genesee and Upshur (2001) describe classroom based assessments and how they are different for ELLs. They explain how classroom observation can allow a teacher the opportunity to see if a student has incorporated a specific language learning strategy. Anecdotal records can also be used in the same way. Harp defines anecdotal records as “the written records a teacher keeps of his or her observations of children” (1991, p. 19). The use of anecdotal records with ELLs would focus on language learning strategies. Checklists and rating scales, like classroom observations and anecdotal records, also can document a student’s successful use of a strategy.

Genesee and Usher describe a portfolio as a “purposeful collection of students’ work that demonstrates to students and others their efforts, progress, and achievements in given areas” (1996, p. 99). They explain that “second language portfolios can have a very specific focus, such as writing, or a broad focus that includes examples of all aspects of language development” (p. 99).

An assessment that involves elements of portfolio, observation and record keeping is the Primary Language Record (PLR). The PLR, which was developed in England, is an effective tool for documenting the development of a language minority student’s language growth in English and other languages. Harp explains the purposes of the PLR:

The PLR is developed to meet three main purposes for record keeping: to inform and guide other teachers who do not yet know the child; to inform teachers who work with the child about the child’s work; and to provide parents with information and assessment about the child’s progress. (1996, p. 285)

The PLR has been adapted in the United States for school districts. It was adopted by the Bureau of Indian Affairs (BIA) to be used in BIA schools. It is simply
called the Learning Record. Part of the BIA adoption included ongoing professional development on the use of the Learning Record.

Involving students in assessment through self-assessments, interviews and conferencing are all effective classroom methods. Student conferences can be used with portfolios or they can be used more widely.

Conferences can be used more widely as part of evaluation, and generally take the form of a conversation or discussion between teachers and students about school work. Conferences can include individual students, several students, or even the whole class; they can be conversations about completed work (as in the case of portfolio conferences) or about work in progress (for example, during a reading or writing activity); and they often focus on activities the teacher has set up expressly to observe and discuss. (Genesee & Usher, 1996, p. 109)

Authentic classroom assessments that are used effectively with ELLs include interactive journals, questionnaires and interviews. Genesee and Usher explain that journals as written conversations between students and teachers can be effective because sometimes students can be more expressive on paper.

Some second language students may feel reluctant to express themselves openly during regular classroom activities for linguistic or cultural reasons. For second language learners who are being educated in the second language, keeping a journal can help them express and deal with issues related to integration into mainstream classes. (Genesee & Usher, 1996, p. 120)

There are many classroom based authentic assessments for reading that are appropriate for ELLs. Pierce explains that “because ESL/bilingual students come from such a wide variety of backgrounds, their prior knowledge about reading and their experiences with reading vary” (2001, p. 65). Classroom based assessments that are available to a teacher immediately can assist in developing appropriate instructional programs for these students.
Many mainstream authentic reading assessments, such as running record, which involves the keeping a record of a student’s reading, or miscue analysis, which involves an analysis of the miscues a reader makes are effective with ELLs. Hill, Ruptic and Norwick explain that miscues are something said or read in place of the printed text and that “we can learn a great deal by analyzing the mistakes or miscues that a reader makes” (1998, p. 100). Pierce recommends the assessment of pre-reading, during reading and post reading strategies for ELLs as a means of understanding how much they are comprehending the material.

By using routine instructional activities for reading and to evaluate English-language learners’ use of reading comprehension strategies before, during and after reading, we can help these students actively apply reading strategies and become independent learners. (2001, p. 81)

The area of authentic assessment of oral language for second language learners has developed tools and protocols that differ more from mainstream assessments than others. O’Malley and Pierce explain that “one of the major responsibilities of any teacher working with English language learners (ELLs) is to enable students to communicate effectively through oral language” (1996, p. 57). An authentic oral language assessment that was developed by the California State Department of Education in the 1980s and developed widespread use is the Student Oral Language Observation Matrix (SOLOM).

The SOLOM focuses your attention on five oral language traits: comprehension, fluency, vocabulary, grammar, and pronunciation. Checklists can be constructed to include any one of a variety of oral language behaviors that you wish to document, including particular grammatical structures, vocabulary, conversational interactions and presentation skills. (Peregoy & Boyle, 2008, pp.137 – 138)

The SOLOM, as other authentic assessments, involves teacher judgment. It can be used in an ongoing formative approach, allowing the teacher to capture information on
oral language on a regular basis, or as a summative assessment, providing baseline information for initial identification or program exiting.

Current Legislative Requirements for ELLs

The current authorization of ESEA, which is the No Child Left Behind Act (NCLB) emphasizes a standards-based accountability system that requires that states set standards and assess against those standards. States must disaggregate this assessment data by student groups that have historically underachieved and report on the progress of all groups. ELLs are included in the general standards and assessment requirements and are one of the “sub-groups” that must be reported separately.

Rivera and Collum explain the increased attention on the needs of ELLs due to the requirements to include them in large scale assessment and accountability:

Over the past 10 years, accounting for the academic achievement of English language learners (ELLs) through standards-based state assessment has become a major national priority. Educators and policymakers have come to acknowledge that ELLs must be enabled to participate meaningfully in statewide assessments and that their performance on these tests must be made publicly available so that their academic progress can be compared to that of other student groups. Including ELLs in state assessments and holding states, districts, and schools accountable for the academic progress of these students, it is believed, will ensure that the needs of these students will be made evident and that educators can respond more appropriately to the instructional needs of this growing population of students. (2006, p. xxxiii)

ELLs also have a separate accountability system for English language proficiency. Table 1 and Figure 2 from USDE demonstrate the accountability system for ELLs. Table 1 shows how ELLs are included in the accountability system in which all students participate. This includes instruction based on state standards and annual assessment of student progress in achievement goals set by the state.
Table 1. Standards, Assessments and Annual Measurable Objective Requirements in NCLB.

<table>
<thead>
<tr>
<th>STANDARDS</th>
<th>ASSESSMENTS</th>
<th>ANNUAL MEASURABLE OBJECTIVES</th>
</tr>
</thead>
<tbody>
<tr>
<td>For LEP Students Only</td>
<td>English Language Proficiency Standards</td>
<td>English Language Proficiency Assessments</td>
</tr>
<tr>
<td>For ALL Students (including LEP Students)</td>
<td>Academic Content Standards (reading/language arts, mathematics, science)</td>
<td>Academic Assessments in reading/language arts, mathematics, and science</td>
</tr>
<tr>
<td></td>
<td>Academic Achievement Standards (reading/language arts, mathematics, science)</td>
<td>Annual Measurable Objectives in reading/language arts and mathematics</td>
</tr>
</tbody>
</table>

ELLs also have another accountability system that includes English language proficiency standards, English language proficiency assessments, and goals for progress in English language proficiency. Figure 2 shows how the two accountability systems overlap for ELLs.

Specifically NCLB requires states to:

- Include ELLs in annual state content assessments of reading/language arts, math and science;
- Disaggregate assessment data to show progress of ELLs in meeting goals;
- Annually assess English language proficiency in the areas of reading, writing, speaking and listening; and
- Report progress in meeting goals of attainment and progress for school districts receiving Title III funds.
USDE has continued to refine the guidance provided on English language proficiency standards and assessments. In a document inviting comment on their proposed framework for high quality English language proficiency standards and assessments, USDE explains the purpose of the English language proficiency standards and assessments:

English language proficiency (ELP) standards are meant to guide the design and administration of local language education programs. ELP assessments are a primary tool for identifying ELLs and measuring students’ progress in their English language acquisition. Such assessments are also used to hold States and local education agencies (LEAs) accountable under Title III for helping ELLs attain English language skills and knowledge, including the language necessary for achievement in the core academic content area. (2007, p. 3)

The greatest impact of Title III and NCLB is its emphasis on assessment and accountability for ELLs. The requirements for state standards, large scale English language proficiency assessment and goals for making progress in English were radical
changes from the past in which compliance with regulations related to ELLs was limited to Office of Civil Rights reviews. Few states had English language proficiency standards or English language proficiency assessments that were aligned with their standards.

Wolf et al. (2008a) explain that there were no available tests for states to use for large scale standards based English language proficiency assessment when NCLB was passed. Likewise, they found varying policies and definitions of proficiency among states.

Funding and ELL Programs

Despite the regulations mandating school districts to provide appropriate educational programs for language minority students, funding has not followed. The responsibility was placed on states and school districts to fund the ESL or bilingual education teachers required by law and alternate language services. Some states and school districts have complied with the need to appropriate funds for services. Some have not. The state of Arizona is currently fighting a lawsuit related to the state’s responsibility to fund language support programs for ELLs (Arizona Education Association website, History of Flores vs. Arizona). This case will be going to the Supreme Court, according to an Arizona Department of Education legislative assistant (J. Jones, personal communication, August 11, 2008).

Along with the requirements for standards and assessments, Title III provides formula funding for states to disseminate to school districts. This move to formula funding was a shift from the previous discretionary grants under Title VII which were very competitive. Unfortunately, the funding under Title III is so limited that, like Title VII, it does not necessarily have much impact.
Funding is addressed as an issue in a review of the history of educational policy, services and assessment for ELLs because funding or lack of funding is an implicit issue in any policy. Kozol, in his writing (2008, Jonothon Kozol website) and speaking engagements has addressed the issues of inequities in educational systems and how funding plays a major role in determining whether students get appropriate services. Though there have been court decisions and legislation mandating educational support for ELLs in the last 40 years, the necessary funding to ensure such requirements are met has not materialized.

Standards Based Education

The assessments required by federal legislation today are based on standards. The standards movement began in the 1980s with an emphasis on setting standards for student learning. Lachat explains that the purpose of the standards based movement in education is “holding all students to high academic standards” in an effort “to improve schools and ensure that no child is left behind in the journey toward the American dream” (2004, p. 1). She explains that:

The standards model is based on several important assumptions: that educators can define standards for what is most important for students to know and be able to do in today’s society; that most students will be able to achieve the standards; that student performance may differ in demonstrating proficiency but will still reflect the defined standards; and that standards will allow for fair and consistent assessment of diverse student performances (Taylor, 1994). Two types of standards provide the foundation for standard-based curriculum instruction and assessment: content standards and performance standards. (Lachat, 2004, p. 3)

The emphasis on standards “represent a different way of thinking about large-scale assessment, and they are playing a central role in school reform” (Lachat, 2004, p. 10). The reason for this is that “student learning is being measured against publicly
defined standards, and performance-based assessment methods are being used to measure student proficiencies" (2004, p. 10). Lachat asserts that important features of standards-based assessments are:

- They focus attention on what is most important to learn.
- They compare students to a standard of proficiency, not to other students.
- They are linked to curriculum and instruction.
- They are intended to establish accountability, as well as stimulate improvement. (Lachat, 2004, p. 10)

Lachat sees the purpose of the assessment of standards as radically different from the testing in that it shifts the "emphasis away from ranking students against test norms to an emphasis on improving student learning" (2004, p. 12). This change, according to Lachat, is particularly important for minority students, including ELLs, because "schools are expected to ensure that all students achieve publicly defined standards of learning" (2004, p. 17).

As stated, NCLB also requires specific standard for English language proficiency. In explaining the purpose of their English language proficiency (ELP) standards the World-Class Instruction Design and Assessment (WIDA) consortium state that the standards are based on theory and research. They also include elements of social and academic English.

The development of WIDA's ELP standards has been in response to recent educational change brought about through theory, research and legislation. First, the vision of language proficiency has expanded to encompass both social contexts associated with language acquisition and academic contexts tied to schooling in general, and particularly to standards, curriculum and instruction. (Cranley, Gottlieb, & Oliver, 2007, p. RG-6)
Assessment and Accountability

Assessment, by itself, is not necessarily controversial. Most would agree with Cobb, when she states that “assessment is a critical component of effective teaching and learning” (2005, p. 20). Gottlieb stresses the empowering aspect of assessment for ELLs when she states that “If reliable, valid, and fair for our students, assessment can be the bridge to educational equity” (2006, p. 1). In fact the rationale behind educational accountability systems is often student achievement, success and equity.

The National Council of La Raza (NCLR), an advocacy group for Latino students, explains that the “educational achievement and attainment gaps between Latinos and other U. S. students remain wide. The gaps are even larger with respect to ELL children and their White and African American counterparts” (Lazarin, M. 2006, p. 2). NCLR believes that NCLB “provides opportunities to narrow these gaps by holding schools accountable for improving academic achievement among all groups, including Latino and ELL students” (2006, p. 2). NCLR commends NCLB for this accountability which had not been part of the previous iterations of the federal legislation. Because states and school districts were not held accountable for these students “ELLs were among the most likely students to be exempted from state accountability systems” (Lazarin, 2006, p. 2).

Not everyone agrees that the accountability system under NCLB is equitable. Crane, Rabinowitz and Zimmerman discuss the role of the NCLB accountability system and the fact that many see it as disciplinary rather than helpful because of the fact that it is externally imposed. They advocate for a locally developed accountability system. “Local accountability plans can supplement state plans that may not acknowledge or
reward incremental growth or improvements in areas important to the local committee” (2004, p. 4).

Those that criticize current accountability systems based on standards and assessments stress the fact that students are individuals. Carini, who has developed a form of assessment that emphasizes observation and affirmation of student’s unique skills and abilities, criticizes the standards and assessment reform movement.

Aimed at generalized solutions, applicable across all schools, models and systems miss the point: the point of human differences and human complexity. It is after all people – and most vigorously, children and youth – who learn, who make sense and meaning of the world, and they simply don’t all do it the same way. What ever (sic) the model, children tend to fall through these technological nets. The more refined and totalizing the model, the more fall through. Human complexity, the complexities of learning, the complexities of teaching, resist systemization. (2001, p. 9)

Ohanian also emphasizes that students are all different. She states that a standardized curriculum “gives nonstandard students no place to go” (1999, p. 2). An education should allow for these differences and allow students who have different talents to explore those talents. Ohanian uses the term “standardaristos” in referring to those that advocate for standards in education. She criticizes a standards-based assessment and accountability system based on content standards in limited areas.

Standardaristos ignore the fact that we need our students to grow up to become chefs, plumbers, child-care workers, musicians, and poets as well as engineers and certified public accountants. Most important, (sic) we need our students to grow up to become parents who nurture their children. (1999, p. 3)

Ohanian emphasizes the important role of the teacher in discerning needs of students and providing a differentiated curriculum. “I’ve said it plenty of times before and I will say it again: The really scary thing about teaching is that we teachers,
particularly those of us in elementary school, teach who we are. We are the curriculum” (1999, p. 9).

Graves echoes the voices of Carini and Ohanian when he states “Testing is not teaching” (2002, p. 33). Graves believes that “instead of preparing children for tests, teachers need to teach the skills that will, in fact, make them better readers” (2002, p. 33). He explains some of the difficulties of depending on testing to assess student knowledge. Most large scale tests are scored by testing companies that use computers to score masses of tests in a short amount of time. “Unfortunately, our quick-scoring computers can’t handle responses that demand written thought or discern which students can initiate and pursue a long-term project or even read books” (Graves, 2002, p. 32).

Graves addresses accountability systems that use testing as measurement. He supports the concept of improvement and accountability. “There’s hardly an educator I know who doesn’t want to document improvement” (2002, p. 35). He is concerned about a system that uses numbers as educational accountability. “Whenever someone applies statistics to human growth and development, there ought to be massive rebellion” (Grave, 2002, p. 36). He emphasizes the role of teachers in stating “Good teachers know that methods must be based on the needs of the particular student and that choosing a method is an art based on professional experience and a knowledge of the child’s interests, abilities and desires” (2002, p. 37).

Current Practices in English Language Proficiency Assessment

NCLB legislation is viewed as both supported the accountability requirements in OCR regulations and contradicted them. Some see the emphasis on proficiency in English and requirements for inclusion in academic achievement assessments and
progress goals as conflicting with student rights for programs that support their specific needs. "One problem is that this law contradicts the spirit of Lau v. Nichols," Krashen and Crawford state. "That is, it fails to recognize the unique situation of ELLs, requiring them to meet the same targets for 'adequate yearly progress' and, in most cases, to take the same standardized tests as English-background students" (2007, p. 56).

Others see the emphasis on ELLs and the policies that hold states and school districts accountable for their progress as consistent with Civil Rights issues. Most would agree that the assessment requirements were demanding. Since the requirements, which included state English language proficiency standards and large scale English language proficiency tests, had not been in place before, no state was prepared.

Abedi reviewed the current status and practice in the area of English language proficiency assessment in the nation in for the purpose of providing information on the existing ELP assessments and discuss the national efforts in developing new ELP assessments based on the criteria required by NCLB Title. He states that "These new mandates have generated significant challenges for states with respect to standards and test development; test validity; and accountability policy development and implementation (GAO, 2006; Abedi, 2004; Crawford, 2002). (2007, p. 5).

Porter and Vega, providing an overview of tests in Abedi's report, explain that:

Over the years, many formal and informal assessments have been developed and used for the purposes of measuring English language proficiency of students whose home language is not English. Many of these assessments, however, do not meet the requirements specified in Title III of the NCLB Act. (2007, p. 93)

They state that in response to NCLB, states developed English language proficiency standards and assessments aligned to the standards. A number of states
received grant funding from USDE and formed consortia to develop the assessments. Other states developed their assessments with the assistance of outside test developers and several states opted to use commercially available assessments, which are either ‘off the shelf’ or augmented or assigned versions of assessments (2007, p. 93).

Several of the largest consortia that received grants from USDE to develop assessments included the State Collaborative on Assessment and Student Standards for LEP Students (LEP-SCASS), the Mountain West Assessment Consortium (MWAC), the World-Class Instructional Design and Assessment (WIDA) Consortium, and the Pennsylvania Enhanced Assessment Group (PA EAG).

These consortia were challenged by many issues in their development process, including limited funding, short timelines, lack of direction and leadership, politics and other issues inherent to group organizations. The projects were minimally funded through competitive funds that were not extended. The timelines for the grants were not only short, but states were under pressure to implement the legislative requirements. USDE had allowed minimal time for the transition to the new standards, tests and accountability program. Guidance was not forthcoming from USDE on what the new assessments should include. Research lagged and was unavailable for large scale academic English assessments.

In the midst of these obstacles, some of the consortia were not able to maintain their momentum. Others were able to not only produce the necessary products, but develop an organization that allowed for the ongoing research and development that assessment involves in the new environment.
MWAC did not continue as a consortium once the test was developed. Originally eleven states were involved with the development. Five states have continued to use the test, adjusting it for their needs. Only one state, Florida, is using the test developed by PA EAG. The LEP-SCASS consortium, which was sponsored by the Council of Chief State School Officers (CCSSO), has had the organization and support of CCSSO to assist in organization. Eight states continue to use the test developed by the LEP SCASS group.

The WIDA consortium has grown more than any other consortia. Able to keep one step ahead of the development process, Boals, the director of WIDA, provided the necessary leadership that allowed the consortium to evolve and grow and expand. Boals explains that:

In 2002, two consultants at the Wisconsin Department of Public Instruction and a consultant from the Illinois Resource Center outlined a plan for promoting partnership of states that would support the development of a system based on English language proficiency standards and assessments.... The outline became the foundation of a federal enhanced assessment grant proposal that included an initial partnership of three relatively small states, Wisconsin, Delaware, and Arkansas. (Bauman, Boals, Cranley, Gottlieb & Kenyon, 2007, p. 81).

Currently there are 18 members in the consortium (World Class Instructional Design and Assessment Consortium, 2008) and the consortium provides a system of standards and assessments for member states, along with training and professional development. WIDA left the Wisconsin Department of Public Instruction in 2006 and is housed at the Wisconsin Center for Education Research, which is part of the University of Wisconsin. Boals feels that the move to a research center supported the direction of the consortium better than the state education agency (personal communication, June 25, 2008).
California is the only state that had an ELP assessment in place prior to NCLB. The California English Language Development Test was first implemented in 2001. Other large states, such as Texas and New York have developed their own. Many states use an adaptation or augmentation of a commercial test.

The development process for these new tests first involved the establishment of standards if there were none in place. Then the developers established a "blueprint" for the test, developed items, with the assistance of educators in the member states and conducted reviews of the items using many of the process consistent with test item review and analysis. Tests were piloted, manuals developed and training conducted. Porter and Vega reviewed each test used in the nation looking at purpose, scoring procedures, alignment and technical properties. They conclude with a summary of the issues:

After tremendous effort in a short period of time, states and test developers have made progress in complying with NCLB Title III stipulations. It is important that the reliability and validity of these assessments be examined and the tests be refined even further. Assessments need to undergo rigorous analysis on an ongoing basis. Continued partnerships between test developers, states and researchers will raise the psychometric standards for English language proficiency tests. In addition, further efforts must be made to use alignment methods in the test development process to ensure that test are valid and reliable measures of state English language development standards and state-adopted content standards. (2007, p. 176)

Wolf et al. (2008b) notice similar concerns in their review of state English language proficiency tests. They observed that states emphasized academic English in the standards they developed, but found a varying degree in the "specifying the nature of academic English proficiency" (2008b, p. 24). They explain that:

Considering that the majority of states used the ELP tests as a primary source to identify ELL students, determine their level of proficiency, and redesignate their
ELL status, the issues of variation in and comparability of these assessments are notable. Another notable issue is the mismatch in proficiency levels established in the standards and those used in the ELP assessments. While states are establishing the technical adequacy of the use of their new ELP assessments, issues of comparability and alignment in the constructs also need to be addressed in their validation arguments. (Wolf et al., 2008b, p. 24)

Related Research

As discussed in the introduction to Chapter II, I conducted an initial review of the literature to establish a basis for research and I continued to review literature related to my topic as I collected my data. Primarily, the literature cites a need for more research related to ELLs and assessment. Also, available research in the area of ELLs and assessment tends to focus on ELLs in the general content assessments. Research is even more limited in the role of teachers' voices in large-scale accountability systems. Though many proponents cite the need for large-scale systems to include authentic assessments and the voices of those who work with students more closely, there is little research to support that assertion.

*Research Related to English Language Learners and Assessment*

Members of the National Literacy Panel on Language-Minority Children and Youth had difficulty finding sufficient research studies in the area of literacy instruction and language minority students to draw conclusions on research based methods. Both August and Shanahan cautioned that their findings were based on very limited numbers of research studies, when presenting on the results of the Panel at the International Reading Association Conference (presentation, May 2, 2006). In the written summary August affirms the importance of the teacher's role in the assessment process and the need for more research:
For placement purposes, there is limited evidence about the effectiveness of teacher judgment in identifying language-minority students who need intensive reading instruction or who might be in danger of dropping out of school. The findings suggest, however, that teacher judgment might be more reliable when teachers can respond thoughtfully to specific criteria rather than to express their opinions spontaneously. Because teacher judgment and assessment play a significant role in the education of language-minority students, additional research needs to explore this assessment tool further. (2006, p. 8)

Bailey addresses the need for research in the area of language minority students in large-scale assessment and accountability systems:

There are major research and policy problems facing the USA and other English-speaking countries with large populations of primary and secondary school students learning English in academic contexts for academic purposes. The problems can be most succinctly articulated as a lack of comprehension information to answer questions about what language demands are placed on school-age children in general, and what language English language learners can realistically be expected to learn and how quickly. While we have a large literature base about English-as-a second language (ESL) this research base cannot answer these questions with sufficient specificity to aid policy-makers and educators faced with the creation of English language development tests and curricula. (Bailey, 2006, p. 3)

Lachat addresses some of the policy and research issues in looking at standards-based instruction and assessment for English Language Learners:

Farr and Trumbull (1997) caution that new assessment practices may have limited utility for English language learners because of the common practice of getting reforms in place for ‘the majority’ and then trying to adapt them to ‘special populations,’ often after financial and human resources have been exhausted. (Lachat, 2004, p. 58)

Lachat continues on to state that “addressing the needs of English language learners as an afterthought makes it more difficult to develop assessments that are inclusive, valid, and reliable for this population” (2004, p. 58). Lachat supports the role of ESL and bilingual teachers in the development of assessment and accountability systems for English language learners.
O’Malley and Pierce discuss the lack of research and information on authentic assessment for English Language Learners:

While there has been a high degree of interest in authentic assessment in general education, as evidenced by the number of articles and books appearing on the topic, there are relatively few articles and monographs on alternate assessment with language minority students (students who speak a language other than English as their first language and/or come from an environment where a language other than English is spoken). The more general articles and books on alternative assessment, while often useful, do not focus on the specific needs of language minority students and often fail to provide specific examples that teachers can use in classrooms. (1996, p. 2-3)

More studies are available in the area of general education and teacher developed assessment tools. These studies document the importance of using classroom-based assessment information in making decisions on student learning and growth.

Black and others explored the formative types of assessments used by educators in the classroom on a regular basis. They looked at several questions that addressed whether formative and classroom based assessments were helpful (Black, Harrison, Lee, Marshall, & William, 2004).

We believe that the three sections below establish a strong case that governments, their agencies, school authorities, and the teaching profession should study very carefully whether they are seriously interested in raising standards in education. However, we also acknowledge widespread evidence that fundamental change in education can be achieved only slowly -- through programs of professional development that build on existing good practice. Thus we do not conclude that formative assessment is yet another "magic bullet" for education. The issues involved are too complex and too closely linked to both the difficulties of classroom practice and the beliefs that drive public policy. (Black, Harrison, Lee, Marshall, & William, 2004, pp. 9-11)

Bailey has continued her exploration of academic English language proficiency in compiling a number of research studies that explored the area. Though she concurs that research is limited (2007, p. 11), she found a number of studies in the area. Butler,
Stevens and Castellon found in reviewing several research studies that the studies “verified long-held beliefs that traditional language tests may not assess the full range of English necessary for students to handle the material on standardized content tests” (2007, p. 46). Bailey, Butler, Stevens and Lord explored the nature of academic language that students encounter in the K – 12 classroom in order to better define the construct of academic English language proficiency (2007). They found that:

There was great variability in the degree to which teachers held students accountable for verbalization of their knowledge, with only some teachers for example, requiring students to provide fully elaborated explanations for their scientific claims. This source of variation in teacher discourse style has implications for student learning and assessment… (2007, p. 148)

Menken reviewed the use of standardized tests with ELLs (2008). Though she focused on the inclusion of ELLs in the academic content tests, the research is significant in that she looked at the policy behind testing. “To date, little research has been devoted to the inclusion of ELLs in high-stakes testing in the United States, particularly with regard to language policy implications of this practice” (Menken, 2008, p. 50). “The reality is that tests offer a highly potent and expedient method for changing school curricula and classroom practices, particularly when the stakes attached to them are high” (2008, p. 8). Her primary interest is in the less explored area of how testing impacts the instructional practices and the learning experiences of ELLs, and analyzing the implications of this through the lens of language policy (2008, p. 9).

Menken found several results of high-stakes testing that could be considered unintended, but contributed to an implicit language policy. Several of her findings relate to those other researchers have addressed, such as the fact that teachers tend to teach to the test and simplify their curriculum. Because of the emphasis on English, she
determined that "high-stakes testing in English serves to increase the already high status of this language in the United States, and is linked to the marginalization of ELL students" (2008, p. 111). Menken discussed that the provision allowing some students who speak languages in which there are significant numbers are allowed to take achievement tests in their own language, while others are not is inequitable. This creates a system in which languages are given status.

Testing has become de facto language policy, greatly impacting language education, resulting in the standardization of test languages and the creation of linguistic hierarchies when some minority languages are officially recognized through test translations and others are not, and the dominance of English is reinforced. (Menken, 2008, p. 179)

Research Related to Assessment and Teacher Impact

An area of research related to my study, but limited in information, is the area of teacher perceptions and attitudes. Several studies have been conducted on general assessment and teacher impact. I reviewed several studies that specifically addressed ESL teacher and identity and voice.

Hamilton reviewed research on assessment as a policy tool and found that the impact of large-scale tests was mixed (2003). There was a positive effect found on student achievement (2003, p. 33). He found that the unanticipated consequences of high-stakes testing included the reallocation of teacher efforts to content that is tested (2003, p. 33) and deemphasizing subjects that are not tested. Hamilton did find that a number of studies attested to the pressure high-stakes testing places on teachers and school districts and the reduced morale that resulted (2003, p. 38).

A study conducted by Taylor, Shepard, Kinner and Rosenthal on teachers’ perspectives on high-stakes testing in Colorado found similar issues (2003). Though the
standards-based reform had caused substantial changes to the curriculum and teaching practices and instruction became more rigorous, there were some negative consequences. Along with the decreased attention to subjects not tested, it was found that there was a decrease in faculty morale (2003, p. 55). Referring to the reports that document the school district progress towards meeting state standards Taylor, et al. explain that teachers are discouraged:

Teach...public confidence in education. They also believe that the consequences following from Report Cards will not affect poorly qualified teachers but will instead cause more well-qualified teachers to leave the profession. Some teachers self-reported that they might retire early because “teaching just isn’t as much fun anymore.” (2003, p. 56)

Nordmeyer focuses on the needs of ESL teachers for support in meeting the needs of ELLs in a high-stakes testing environment. He states that ESL teachers are essential in helping ELLs to face the twin challenges of achieving in content areas and developing academic English proficiency (2008, p. 36). ESL teachers need to collaborate with mainstream teachers and both teachers need professional development.

Tsui studied identity formation of an EFL teacher in Hong Kong looking at her professional identity. Tsui found “identity formation is highly complex” (2007, p. 678), and being part of a community is part of the identify process. Tsui explains that “participation as well as nonparticipation in negotiating meanings is shaped by power relationships among members of a community” (2007, p. 678).

An area of research related to teacher perceptions, attitudes and voice in assessment and accountability is the area of students and communities. The impact of assessment on children and families can be viewed in a number of ways. One perspective
is that research that documents the success or failure of large scale assessment represents the student and community voice. Another perspective advocates for different ways of assessing.

Some researchers and policy makers document the success of the current large scale accountability systems. As discussed, the National Council of La Raza supports an accountability system that includes language minority students in the testing program and reports the results separately (2006). Bowe, Cronin, Kingsbury & McCall in looking at student achievement scores found that in a 2005 study that NCLB was both successful and not successful. They found that student math and reading scores have improved. Students across the nation are making more progress in meeting each states' standards. On the other hand, when student scores were compared on a growth model they have decreased. A growth model would look at the growth of individual students instead of the performance of each grade as NCLB requires. Bowe et al. found that the scores of ethnic minorities have not increased under NCLB (2005).

Other researchers advocate for other ways of teaching and learning and assessing. Fayden's (2005) work in how children learn represents the voice of the students and community. Fayden studied a small group of students who were categorized as language minority and found that her students had rich linguistic backgrounds and ways of learning. She explains some of the problems with testing. "The point I want to make is this: The use of inadequate testing can cause inadvertent racism because it leads educators to view their charges as having shortages" (2005, p 23). She explains how she accepted the idea the students were deprived and deficit in language when she first started teaching in a small reservation school in the Southwest.
For several years I accepted the assumption of these children being delayed. The entry and continuing levels were proscribed in the tests that we administered to the children, and our students clearly appeared to be at an impasse. Yet I was puzzled when I was told that the children needed numerous oral language lessons because their language fell short (compared with the mainstream). My students had abundant vocabularies and seemed to communicate their needs and thoughts in competent ways. The very social nature of life in their village seemed to give the children an affable manner, portrayed by ease and friendliness in conversing with others. (2005, p. 17)

Fayden’s research found that instruction and assessment in a social constructivist model allows educators to see the strengths many children bring to the classroom and is empowering, rather than focusing on deficits.

McCarty (2005) also addresses the voice of the student and community in educational accountability. She explains that the emphasis on standard American English and academic testing diminishes the rich heritage students bring to the classroom.

In the United States, the forces for standardization have reached new and alarming proportions, with federal and state policies mandating scripted, uniform, remedial reading programs for students identified as ‘at risk,’ ‘limited English proficient,’ and ‘deficient in reading skills,’ The intended targets of these policies are clear. The 2001 federal No Child Left Behind Act, for example calls specifically for implementing these pedagogies with ‘underachieving’ American Indian, African American, and Hispanic Students. (2005, p. xv)


Summary

Learning and language theory are important issues to review in looking at assessment in general and specifically English language proficiency assessment. Language learning theory and definitions of language proficiency continue to be controversial. Expectations for language proficiency increase as the demands for literacy
in our world increase. Social constructivist theory best exemplifies the complexity of language learning.

Research studies addressing assessment, the impact of assessment and specifically English language proficiency assessment are limited. Research studies, in general, on educational practices with language minority students, are limited. Research studies involving the voice of teachers in large scale accountability systems are also limited. Studies which address the voice and knowledge of ESL and bilingual teachers in large-scale and high-stakes assessment practices are even more limited.

There is more information available on the use of authentic and formative assessment with general education. This research is supportive of the important role this type of assessment can place in documenting student learning. More research is needed, though. The majority of authors addressing educational issues related to language minority students cite the need for more research on instructional and assessment practices.
CHAPTER III

PROCEDURES

A description of the research design and procedures used in my study is covered in Chapter III. The research design and the assumptions underlying the design will be first addressed, with a description of the research location and participants following. Details of the data collection methods and procedures come next, with specific descriptions of each tool and activity. A description of the data analysis procedures follows the collection methodology section. The discussion of the data analysis specifically addresses how the data for each tool and activity was analyzed, including samples of instruments and data analysis process to demonstrate my arrival at themes, codes and categories. The researcher role is finally discussed, with an explanation of verification, trustworthiness, researcher bias and limitations.

Research Questions

The foundation of the research design is the research problem and related questions. The questions determine the type of data collection tools and analysis.

The questions include:

1. What assessments are English as a Second Language (ESL), English Language Learner (ELL), and bilingual teachers using to document the development of English language proficiency of English Language Learners?

2. What do the teachers say the assessments reveal about the students?
3. What is the relationship between teacher developed assessment information and large-scale assessment data and accountability?

Design of Study

The research design for the study is a mixed method study, based on a naturalist constructivist paradigm. I use both terms “constructivist” and “naturalist” to define my overall research model, which reflects the world view and theory on which the research is based. The terms “qualitative” and “quantitative” are used as terms for methods. Because most of my data analysis involves theory building, inductive reasoning and exploration, which are consistent with naturalistic research, the overall design supports a naturalistic or constructivist paradigm.

The literature uses both the terms naturalism and constructivism in reference to research which emphasizes inductive reasoning, exploration, theory building and the emergence of new meanings in the research environment. Naturalism is typically aligned with qualitative research methods which are “used when observing and interpreting reality with the aim of developing a theory that will explain what was experienced” (Benz & Newman, 1998, p. 3).

In contrast to a positivistic view of the world, which accepts a common reality on which people can agree (Benz & Newman, 1998, p. 2), the naturalistic perspective allows for multiple realities, multiple interpretations that may be considered valid. Gherardi and Turner reference the work of Glaser and Strauss (1967) in explaining that naturalistic research is more concerned with discovery, rather than verification (2002, p. 90). Glesne further explains the philosophy supporting qualitative research methods, making the connection with constructivism.
Positivists assume a fixed, measurable reality exists external to people. In contrast, qualitative methods are generally supported by the interpretivist (also referred to as constructivist) paradigm, which portrays a world in which reality is socially constructed, complex, and ever changing. (Glesne, 1999, p. 5)

Typically qualitative research methods are consistent with a constructivist paradigm. I use both qualitative and quantitative research methods. Benz and Newman (1998) argue that the researcher chooses methods based on the need for specific data. I chose my methods based on the type of questions asked and the data I wanted. “The decision about what data to collect, as well as what to do with those data after they are collected, should be dictated by the research question” (Benz & Newman, 1998, p. 15). Time and scheduling also entered into the decision-making on data collection. The single quantitative tool, a questionnaire, was purposely used at the beginning. Using the questionnaire as the first data collection activity allowed for some baseline data that was easily collected.

Typically a mixed method design “places more emphasis on one type of data than the other in the research and the written report” (Creswell, 2002, p. 569). The emphasis in my study is on the qualitative methods, analysis and viewpoint, since they are more supportive of a naturalistic research model. Whereas the single quantitative activity and tool used in my study has a purpose and value, the data gained the qualitative methods have a stronger role in data analysis and interpretation.

Figure 3 provides a visual overview of the research design and data collection plan. The research questions, which are listed in the upper left hand corner, serve to guide the study and the design. The phases of research activities are represented by red rectangles and include:
Research Questions
1. What assessments are ESL, ESOL Bilingual teachers using to document the development of English language proficiency of English Language Learners?
2. What do the teachers say the assessments reveal about the students?
3. What is the relationship between teacher developed assessment information and large scale assessment data and accountability?

A: Review Theory & Literature
Theoretical basis of study:
constructivist & socio-cultural theory

B: Collect Data

1. Questionnaire data
2. First Interview data
3. Focus group data
4. Second Interview data

C: Final Step:
Summarize data for new meanings & theory

Participants:
ESL & Bilingual Teachers

Sub-steps after each data collection step
1. Analyze data
2. Review literature & theory with data collected
3. Incorporate new meanings

Background landscape
Increased emphasis on accountability
Large scale assessment mandates
Increased numbers of multicultural students
Teachers' voices lost in broadcast of federal mandates
Greater need for citizens with high levels of English literacy and proficiency

Figure 3. Research Design.
A: Review of theory and literature,
B: Collect data, and
C: Final steps, summarize data for new meanings and theory.

The theoretical basis of the study is constructivist and socio-cultural learning theory which is represented by the yellow oval, symbolic for light that continues to illuminate the data collection activities. The participants in the study are also represented by an oval, similar to the oval that is used to provide the label for the theoretical basis. The participants, like theory, provide information for the research to be conducted. They are placed centrally in the diagram, affirming the importance of the research participants and the fact that all the activities revolve around them.

Data collection activities, which are part of step B, are labeled in light blue. The first activity is labeled in a rectangle shape, signifying that it is a quantitative activity. The others are ovals, representing their qualitative nature. Lines are used to connect the activities with the theory, participants and steps. Lines also radiate from the theory oval, representing the concept of light shining.

Theory informs all activities in the research. This concept is represented by the lines symbolizing the interconnections between the research activities, steps and theory are overlapping. The overlapping quality also is representative of the idea that the research is somewhat cyclical.

The data collection activities are placed on the outer edges of the diagram and circle around the theory and participant symbols. Their placement in the outer parameters symbolize though they are influenced by the other aspects of the research.
design. The participants, illuminated by theoretical basis of the study, are the heart of the study and give “life” and meaning to the data collection activities.

At the bottom of the diagram is the “background landscape” of the research study, symbolized by diagonal green forms labeled with different statements. This background landscape represents the current environment in which the participants work and the data is collected. It includes educational requirements, activities, viewpoints and demographics. Data is collected, analyzed and interpreted, mindful of this landscape. Ultimately, this landscape is the reality which influences perceptions and interpretations. When analyzing qualitative data it is important to consider what all of it means to the participants. Thus the background landscape is a crucial aspect in data collection, analysis and integral to the research design.

Missing in the diagram is the researcher, herself. Though the researcher would be expected to leave biases, expectations and belief systems behind in entering the research field, thus, eliminating a need to be represented in the design; she is, in fact, a significant part of the process. The researcher makes decisions along the way of the project, rephrasing interview questions, exploring connecting paths in interview discussions, exploring the same areas in literature reviews, and finally discerning patterns from the data in the final analysis. “Data do not develop theory; people do” (Benz & Newman, 1998, p. 17). To add the researcher, though, would require that the diagram to move to a more multi-dimensional model and include motion since she would need to be a traveler in the diagram.

The interviewer-traveler wanders through the landscape and enters into conversations with the people encountered. The traveler explores the many domains of the country, as unknown territory or with maps, roaming freely
around the territory. The interviewer wanders along with the local inhabitants, asks questions that lead the subjects to tell their own stories of their lived world, and converses with them in the original Latin meaning of conversation as “wandering together with.” (Kvale, 1996, p. 4)

In place of a diagram that uses motion I would ask the reader to use his or her imagination and travel with me through my research project. This journey will take you through over hills and highways, figuratively and literally, and will take you into conversations with teachers. Join with me in hearing these conversations, along with the other responses from the participants. Listen to their voices and the information and knowledge they share about students and assessment.

The specific research tools used involve a questionnaire, interviews and focus groups. I used a questionnaire as the first data collection instrument. Most of the data from this instrument, other than the open-ended questions is quantifiable. Appendix B provides a copy of the questionnaire instrument. This quantitative tool was used, as explained, because its purpose was to provide some initial concrete questions which could be analyzed easily to inform the next step. The other three data collection activities, including the two interviews and focus groups were analyzed using qualitative research techniques. Each data collection activity informed the next activity in an effort to better understand the issues, confirm or not confirm data collected previously.

The data collection tools and methods were chosen because of their usefulness and the type of data they would collect. They also serve as validation. Creswell explains that a mixed method design can be used for triangulation purposes in that both quantitative and qualitative can be merged and the results can be used to best understand the problem (2002, pp. 564-565). "The researcher gathers both quantitative and
qualitative data, compares results from the analysis of both data, and makes an
interpretation as to whether the results from both data support or contradict each other”

The questionnaire data allowed participants to respond to concrete questions
related to the research topic in a traditional survey format. While this instrument did not
offer an opportunity for the exploration or development of ideas, it yielded data that
could be used for discussion for interviews and focus groups. It also provides a means to
validate the data that came later. The interviews, which supplied the richest data, were
designed to allow participants to share, discuss and explore topics in an effort not only to
validate data, but also create new meanings and understandings. The focus groups
provided another qualitative method of data collection that differed in that it allowed for
the interaction of individuals in a group.

Thus, the design is mixed in a number of ways. Both quantitative and qualitative
methods are used in order to provide activities most supportive of the appropriate data
and analysis. But, also the activities, themselves, are conducted in different media,
including paper questionnaires, face to face interviews and online focus groups. This
variety not only allowed for triangulation of data, but also allowed participants different
opportunities to respond. Some people may be more comfortable writing their thoughts
rather than talking. Others may be more comfortable discussing ideas in a group rather
than in an interview. A variety of methods ensured that these different preferences would
be accommodated. The variety of methods and media also, as discussed, supported
validation efforts in that one could attribute greater trustworthiness to concepts and codes
that surfaced in different activities.
All research activities were approved through the University of North Dakota Institutional Review Board (IRB) prior to the study. IRB policy concerning protocol for human subjects was carefully observed to ensure that there is no risk to participants and confidentiality is observed. A consent form outlining the study was used and is included in Appendix A.

Location

The research study took place in three Midwest states. The study had proposed several states in an effort to include participants from different perspectives. The specific states were chosen for several reasons. All of the states have a mixed ELL population, encompassing different ethnic backgrounds, levels of proficiency and immigrant status. I was also able to include school districts of different enrollment levels in communities that were both urban and rural. Though only three states were used, the distances involved were great. There is over 1000 miles between the school districts of the participant who was the furthest west and those that were in the eastern state. Thus, though I chose states “in my own backyard” so to speak, since they include my home state and neighboring states, the states still reflect a wide diversity of students, teachers, populations and policies.

Table 2 provides demographic information on the states. The table includes information on the total enrollment in the states for students from kindergarten through the 12th grade. Enrollment for ELLs is included, along with the percentage ELLs are of the total enrollment. State requirements are included to provide background on the state commitment to the ELL population.
Table 2. Participating States.

<table>
<thead>
<tr>
<th>State</th>
<th>State I</th>
<th>State II</th>
<th>State III</th>
<th>National</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total K – 12 Enrollment</td>
<td>144,481</td>
<td>103,706</td>
<td>900,347</td>
<td>49,324,849</td>
</tr>
<tr>
<td>Total K – 12 ELL Enrollment</td>
<td>6975</td>
<td>5,737</td>
<td>63,364</td>
<td>5,074,572</td>
</tr>
<tr>
<td>ELL Percent of total enrollment</td>
<td>4.83%</td>
<td>5.53%</td>
<td>7.04%</td>
<td>10.29%</td>
</tr>
<tr>
<td>State requirements for ESL and/or bilingual education</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>State funding for ELL and/or bilingual education</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Meets NCLB requirements for English language proficiency assessment</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Name of NCLB mandated English language proficiency test</td>
<td>State Comprehensive Assessment System <em>(State CAS)</em> for English Language Proficiency</td>
<td>Assessing Comprehension and Communication in English State to State for English Language Learners <em>(ACCESS for ELLs)</em></td>
<td>Test of Emerging Academic English <em>(TEA)</em> &amp; Student Observation Oral Language Matrices <em>(SOLOM)</em></td>
<td></td>
</tr>
</tbody>
</table>

The participating states will be referred to as State I, State II, and State III. State III has the greatest population and largest number of urban communities. The state also has the largest number of multicultural students and ELLs. This state has also had state requirements in place and state funding for ELLs for a number of years. State I has no state requirements for teacher licensure for bilingual, ESL or ELL. Nor does it have any dedicated state funding source other than the federal Title III English Language Acquisition Program. State II has state requirements that school districts provide services and standards for ESL and bilingual education teachers, along with funding.
All states participate in the federal ESEA programs, including Title I and Title III of NCLB, which provide funding for ELLs and requirements for assessment. Since all states and school districts that benefit from any kind of federal funding must be compliant with OCR, the participating states and the school districts in them would be required to provide services and appropriately trained teachers for ELLs, since these students are covered under Civil Rights legislation. They must also follow the same requirements for assessment addressed in NCLB. These requirements, as discussed in Chapter II, include the requirement that ELLs must participate in the statewide content tests along with participating in the statewide English language proficiency tests.

Participants

My research design proposed the selection of nine to 12 teachers to participate in the study. My interest was to include teachers who worked with ELLs, had at least five years of teaching experience and had coursework in the area of ESL and bilingual education. Since only two of the three states had specific licensure requirements for ESL and bilingual education, I was more flexible in this area. I also wanted to include teachers who worked with students from different backgrounds. I made an active effort to ensure I included participants who worked with Native Americans, Hispanics, new immigrants and refugees. My criteria also included teachers in rural and urban areas.

The teachers were recruited through my work contacts. Several of the participants I had known in advance and the others I met through recommendations from colleagues in other states. Nearly all participants were women, which is not be reflective of society as a whole, but is reflective of the field of teaching and particularly teaching ELLs. The only participant who was male, dropped out after the first activity.
Eleven participants initially agreed to participate in the study. One individual dropped out after the first data collection activity. Since my research design and number of participants allowed for some attrition, I used the data he contributed and did not seek another participant. Ten participants completed three activities. Nine completed all four activities. I had one participant who was unable to participate in a focus group. I sent her the transcript so that she could respond.

Table 3 shows the demographics of the participants, including information on their school districts, students, educational level and other relevant information. This table provides an overview of the participants. Three participants came from State I, four came from State II, and four came from State III.

As the table demonstrates, all the participants have significant experience and education. All the participants have at least ten years of teaching experience, with three of them having over 25 years. One teacher is close to retirement. All of the participants have had coursework beyond a Bachelor’s Degree, with five of the participants holding a Master’s and one a Specialist Degree.

All participants work in schools with English Language Learner or Bilingual Education Programs. The participants work in some capacity involving language development within the programs. The majority of participants work with English language development, with two being responsible for heritage language development. The two that worked with the heritage language in a bilingual or home language literacy context, also assisted with English language proficiency assessment, therefore allowing them to be able to participate in the study.
<table>
<thead>
<tr>
<th>Participant (pseudonym)</th>
<th>Grade Level</th>
<th>Teaching Years</th>
<th>Position</th>
<th>Education</th>
<th>Ethnic Background</th>
<th>Enrollment of District</th>
<th>State</th>
<th>Ethnic Population of Students Served</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amy</td>
<td>K-12</td>
<td>25-30</td>
<td>Curriculum Coordinator</td>
<td>Master's</td>
<td>White/ American</td>
<td>1000 - 2500</td>
<td>I</td>
<td>Mostly Native American</td>
</tr>
<tr>
<td>Catherine</td>
<td>K-12</td>
<td>15&lt;</td>
<td>ESL Teacher</td>
<td>Master's</td>
<td>White/ American</td>
<td>5000 - 7500</td>
<td>I</td>
<td>Immigrant, Hispanic, Nat American</td>
</tr>
<tr>
<td>Nona</td>
<td>9-12</td>
<td>15-20</td>
<td>ESL Teacher/ Coordinator</td>
<td>Bachelor's</td>
<td>White/ American</td>
<td>2500 - 5000</td>
<td>I</td>
<td>Immigrant, Hispanic, Nat American</td>
</tr>
<tr>
<td>Iva</td>
<td>K-12</td>
<td>10-15</td>
<td>ESL Teacher</td>
<td>Master's</td>
<td>White/ immigrant</td>
<td>7,500 - 10,000</td>
<td>II</td>
<td>Immigrant, Hispanic, Nat American</td>
</tr>
<tr>
<td>Sharla</td>
<td>K-12</td>
<td>25-30</td>
<td>Bilingual Teacher</td>
<td>Bachelor's</td>
<td>Native American</td>
<td>1 - 1000</td>
<td>II</td>
<td>All Native American</td>
</tr>
<tr>
<td>Viola</td>
<td>K-12</td>
<td>10-15</td>
<td>ELL Program Coordinator</td>
<td>Master's</td>
<td>White/ American</td>
<td>12,500 +</td>
<td>II</td>
<td>Mostly new immigrant</td>
</tr>
<tr>
<td>Wanda</td>
<td>6-12</td>
<td>10-15</td>
<td>ESL Teacher</td>
<td>Bachelor's</td>
<td>White/ American</td>
<td>12,500 +</td>
<td>II</td>
<td>Mostly Native American</td>
</tr>
<tr>
<td>Della</td>
<td>9-12</td>
<td>30+</td>
<td>ELL Program Coordinator</td>
<td>Master's</td>
<td>White/ American</td>
<td>15,000 +</td>
<td>III</td>
<td>Mostly new immigrant</td>
</tr>
<tr>
<td>Don</td>
<td>9 – 12</td>
<td>20+</td>
<td>ESL Teacher</td>
<td>Master's</td>
<td>White/ American</td>
<td>15,000+</td>
<td>III</td>
<td>Mostly new immigrant</td>
</tr>
<tr>
<td>Hettie</td>
<td>K – 12</td>
<td>10-15</td>
<td>ESL Teacher</td>
<td>Bachelor's</td>
<td>White/ American</td>
<td>1000 - 2500</td>
<td>III</td>
<td>Mostly second generation Laotian</td>
</tr>
<tr>
<td>Kim</td>
<td>K-6</td>
<td>15-20</td>
<td>Spanish literacy Teacher</td>
<td>Specialist</td>
<td>White/ American</td>
<td>15,000 +</td>
<td>III</td>
<td>Mostly Hispanic (teaches Spanish)</td>
</tr>
</tbody>
</table>
Most of the participants have direct contact with students on a daily basis. Three of them serve as program coordinators, with one serving as both coordinator and teacher. Those that serve as coordinators continue to have close contact with students and have had recent classroom experience. All participants are responsible for a number of tasks in their roles. The majority of participants serve students at all grade levels, with three individuals are assigned to the secondary level and one is assigned to the elementary level.

The participants include those from small, medium and large school districts. The school districts are located in urban and rural areas. An inner city school district is included. Two school districts are located on Indian reservations. A diverse population of English Language Learners is served by participants in the study from new immigrant to Native American and Hispanic.

The participants are also diverse in their backgrounds. While the majority of them come from backgrounds that were consistent with the majority culture in our nation – white, Anglo-Saxon and Protestant. Two of the participants in the study were from the same cultural background as some of their students. In addition, several of teachers from the majority American culture had lived in other countries and experienced the issues involved with being a person who speaks a minority language. (Note: Individualized information on each participant follows).

Participants were initially recruited in the late fall and early winter of 2007. Data collection activities for the project began in January of 2007 and ended in September 2007. I connected with the teachers through telephone and e-mail and followed through
by sending a letter of consent to them and then the paper questionnaire. This questionnaire is available in Appendix B.

Once I received a signed letter of consent and completed questionnaire, I scheduled the first face to face interview. The first interview took place in March with the last interview taking place in June. The online chat focus groups took place in June and September. The final telephone interviews took place in September. Table 4 provides a list of dates of completion.

Table 4. Data Collection Activities Schedule.

<table>
<thead>
<tr>
<th>Participant</th>
<th>State</th>
<th>Completed Questionnaire</th>
<th>Conducted 1st Interview</th>
<th>Participated in Focus Group</th>
<th>Conducted 2nd Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Don</td>
<td>III</td>
<td>3/2007</td>
<td>Did not complete</td>
<td>Did not complete</td>
<td>Did not complete</td>
</tr>
</tbody>
</table>
Participants From State I

The following descriptions of the participants provide additional information to Table 3. The descriptions serve to further introduce the participants and also give a context to the study and the research findings.

Amy

Amy's school district is located on an Indian reservation that was located a distance from well traveled highways and cities in the state. The building Amy works in is an older building with walls and floors that have been painted and varnished many times over. Amy's office, not far from the classrooms with student voices, is full of materials, posters and educational supplies. She serves as a curriculum coordinator, which also covers the area of assessment in her district. She has worked in a number of educational situations on reservations and off in her state.

Catherine

Catherine works in a mid-sized city in State I. She is one of the ESL teachers for the district. She used to be the only one and traveled extensively throughout the day, but the district has hired another one which has lightened her load a bit. Her classroom is in one of the school buildings that is near the university. It is a small room which might have been a storage room at one time. It has room for a table and it is filled with educational materials, pictures and supplies. Catherine first began her career teaching overseas and continues to spend summers traveling. She organizes a number of multicultural events, along with her teaching.
Nona

Nona works in a school district that is located the furthest west of any of my participants. I did not make it to her school, but she describes the district well. Nona and I conducted our first interview at a conference on multicultural education. Nona teaches and coordinates curriculum and assessment. Her tasks are multiple and varied. Because the district has a diversity of students qualifying as ELL, she works with a variety of students, from the new immigrant refugees, to Native Americans.

Participants From State II

Sharla

The interview with Sharla takes place in her home. She lives in the country, not far from the small reservation school at which she teaches. Sharla is Native American and lives on the land of her ancestors. Sharla, like Kim in State III, has the position of teaching the students' heritage language. Unlike the Spanish that Kim teaches, Sharla teaches Arikara, which is an indigenous Native American language, and has far fewer speakers than Spanish. Sharla incorporates many cultural activities into her teaching.

Iva

Iva is a newcomer to her community. Though she was a teacher in her former country, she came as a refugee and was forced to take any available work until she could get some more education and upgrade her credentials so that she return to her field. Iva not only upgraded her education to get licensed to teach, she had continued on with a master's degree. Iva shares the ELL program responsibilities with another ESL teacher who has been in the position for a longer period of time. They both have heavy caseloads as the numbers of students increase in this district.
Viola

Viola is coordinator of a large ELL program in her state. She is responsible for the administration of the program and works directly with teachers instead of students. Most of the ELLs in Viola’s district come as refugees and have experienced trauma in wars and civil upheavals. Fairly new to the position, Viola has spent years in the classroom and has credentials in a number of educational areas. I have known Viola for some time and have been in the classroom and observed her skills as a teacher.

Wanda

Wanda’s classroom/office, in contrast to many of the others that I interviewed, is a desk in the hallway. Students come and go and I need to adjust my tape recorder so that the background noise isn’t as distracting. She clarifies that her role has changed with the acceptance of a position with a federal grant to a teacher coaching position. Previously she had been a language teacher, ESL teacher aide and ESL teacher. Currently her primary focus is to work with teachers, assisting them in strategies to meet the needs of the students. She also does a little bit of direct work with students and assists with the English language proficiency assessment. The population she serves through the grant is Native American, but, because of her experience with the ELL program, she also has an understanding of the immigrant students.

Participants From State III

Della

Della works in a nondescript structure tucked away in an older part of the large urban city in State III. It is an older building that houses a number of educational programs and had recently been renamed for a well-loved legislator in the state. Many of
the students are older than average high school students who have newly immigrated to the United States. The school acts as both a newcomer center for new immigrants and alternative high school for students who need specialized curricula. Currently serving as a coordinator for the alternative high school for new immigrant students, Della is close to the end of her career. She has not only taught many years in the classroom in her current school district, but has also taught overseas.

Don

Don works with new immigrant students in the same building as Della. His room is active with lots of energy. Don has had many years working with the new immigrant population. He has a busy schedule and we are unable to schedule an interview which eventually leads to his non-participation in the project.

Hettie

Hettie shares many of the characteristics of teachers who work with English Language Learners, in that she has many responsibilities in her role as the district ESL teacher. Her room is a regular sized classroom, unlike many support teachers and is filled with activity centers and educational materials. She has a sink with water running so that students can always have cold water. Hettie has been part of this small community for some time and has worked with the ELL population, assisting with refugee resettlement previous to her work as the ESL teacher in the district. She knows the parents and the relatives of the students. She has been in the district since the program began. There used to be two ESL teachers in the district, but because of budget cuts, there is now only one.
Kim

Kim works in a school building in an older part of an urban area that is difficult to find. The buses lined up outside the building indicate a school. The building is old with long, dark hallways; however, Kim’s classroom is large and bright. It is filled with a variety of books, materials, language and artifacts. Spanish is the dominant language. Kim’s position is to teach Spanish literacy. She does assist in the English language proficiency assessment, which gives her a perspective on assessment, but her primary responsibility is in the child’s home language. Kim, also, has spent working in a Spanish speaking country and has an understanding of living in a different culture which she brings to her teaching situation.

Protecting Confidentiality

All participants signed a Consent Form which had been approved by the University of North Dakota Institutional Review Board (IRB). An example of this form is available in Appendix A. Part of the initial discussion with the participants was an explanation of how the information would be kept confidential. I also reviewed some of this information in the face to face interviews. None of the participants appeared to be concerned about confidentiality. Pseudonyms are used with all participants. A master list of each participant and her assigned pseudonym is maintained and stored in a locked file cabinet in my home. All interviews were recorded with a digital recorder which is stored separately. A working copy of each file is stored on my computer.

Data Collection Methods

The following section will address the data collection methods and activities. As discussed, my research data collection activities for the study were determined by the
research problem and questions. Collection procedures involved four distinct activities; questionnaires, first interviews, online focus groups and second interviews. These procedures will be specifically addressed in separate sections. The activities and methods followed the original design with no major changes. Some small adjustments, involving scheduling, were made in order to facilitate the process of data collection. These changes will be discussed in the sections discussing the methods.

Data was collected, stored, and analyzed through a combination of methods. Electronic media such as e-mail and the Internet assisted in communication and the online chat focus groups. Interviews were recorded through the use of a digital recorder, allowing the files to be stored and accessed through computer files. Software is included with the recorder, Sony™ Digital Voice Editor 2, allowing transcription. A Sony™ Transcribing kit (FS-85USB), including a foot pedal facilitated the transcribing process. A RadioShack™ Wireless Phone Recording Controller allowed the telephone interviews to be recorded digitally, stored and accessed through the computer. The online chat focus groups were conducted online using the “chat” format of a Course Management System (CMS). The University of North Dakota uses Blackboard™ as their CMS and I was able to utilize the format in the same manner an instructor would use the system to hold class discussions.

I used notebooks and computer programs to keep track of data, schedules and time tables. I put all the participant information into a database and noted the date as activities, including data collection and analysis, were completed. Table 4 provides example of a table that was generated from the database that listed all the information on the activities.

118
As Table 4 demonstrates, the collection activities spanned six months. The first questionnaires were returned to me in March and the final telephone interview was conducted in late September. Data from the previous activity was being analyzed as the next activities were being conducted. The initial research design proposed a schedule in which each data activity was completed before and analyzed before the next one was started. Because of distance, time and the busy lives of participants, the activities overlapped to some extent.

Questionnaire

The first data collection activity involved a questionnaire of teacher practices and perceptions related to the assessment of English Language Learners. A copy of the instrument is available in Appendix B. This questionnaire was primarily designed to generate quantitative data that could be analyzed statistically. It was designed on the order of an attitude scale which determines “what an individual believes, perceives, or feels about self, others, and a variety of activities, institutions, and situations” (Airasian & Gay, 1996, p. 156). Some questions were open-ended, allowing for narrative answers.

I reviewed literature on survey and questionnaire research, using a combination of concepts. Fink defines surveys as “information collection methods used to describe, compare, or explain individual and societal knowledge, feelings, values, preferences and behavior” (2006, p. 1). She explains that a questionnaire is an instrument used in a survey (2006, p. 1). The development of strong questions is an important part of a questionnaire (Fowler, 2002). I used a five point scale, which is common in attitude scales. Most of the items used a Likert scale format, but not all. A Likert scale uses a format that
requires participants to respond to a series of statements along a continuum which each response associated with a point value (Airasian & Gay, 1996).

The questionnaire included 38 questions. All but two items used the five point scale. The other two were open ended questions. Participants also had the opportunity to respond with more explanation.

The questionnaire was designed over a period of time. It was originally designed as a survey that would involve a large number of educators in a single state. The questions addressed issues related to the literature on English language proficiency assessment and current practices. Other educators assisted in the development of the questions and it was reviewed by a state advisory committee. When my research focus changed to use interviewing as the primary data collection method the questionnaire was slightly revised. The revised version was informally reviewed by other educators, along with my advisor.

The questionnaire is broken into three parts, including: Part I: General Questions on English Language Proficiency Practices, Part II, Indicators of English Language Proficiency, and Part III: No Child Left Behind (NCLB) and English Language Learners. Part I was also divided into different categories, including Sources of Information, Use of English Language Proficiency Assessment Information, and Types of Assessment Information. Participants had the opportunity to respond by checking a series of responses. For example, participants could check “Strongly Disagree,” “Disagree,” “Neutral,” “Agree” or “Strongly Agree” to the statement: “Information I use to understand my students’ level of English language proficiency is primarily gathered through authentic or classroom based assessment.” The questionnaire also asked for
demographic information questions that provided additional information on the participants. This was used to both verify information that I already had obtained and add additional information. Other than the background information, the responses were intended to be anonymous. I analyzed the data as a whole, without attributing a response to a specific participant.

The questionnaire was used as an informal tool and not intended for wide-scale use. It was intended to be easy to complete. Most of the questions simply required the participants to check a box. Yet, it served to open some thought-provoking issues that could be addressed in more depth in the interviews. Consistent with my interest to start with a "low-key" type of data collection method, the questionnaire was designed in a traditional paper and pencil format. I felt that this would be a familiar method of gathering information that would be comfortable for my participants. I also sent the questionnaire by e-mail. This allowed individuals who were more comfortable with word processing and e-mail to have that option.

The questionnaire was distributed after I had gained the consent of the participants giving them the choice to respond in either manner. Most of the participants returned the questionnaire in the enclosed, stamped envelope I had included. Some questionnaires were returned at the time I conducted the first interview. Though the final analysis of the questionnaire data was not conducted until most of the first interviews had been completed, I was able to provide a summary of responses to generate discussion for the interviews. I also posted the questionnaire summaries on the Blackboard website that I used for the focus group discussions. Consistent with the use of this tool and data to
provide some initial data to inform the rest of the activities, this data was shared with the participants to inform the focus group discussions.

Interviews

The interviews are the heart of my research study and design. They were designed to generate the greatest amount of data and were intentionally conducted in a manner consistent in the literature with qualitative interviewing. Kvale's discussion of the conversational quality of interviewing was a guide for me.

The research interview is based on the conversations of daily life and is a professional conversation. One form of research interview—a semi structured life world interview—... is defined as an interview whose purpose is to obtain descriptions of the life world of the interviewee with respect to interpreting the meaning of the described phenomena. (Kvale, 1996, pp. 5-6)

The interviews were conducted in compliance with protocol that was approved in my IRB. This protocol is available in Appendix C. The protocol followed in the interviews was a combination of informal discussion and focused questions. My interest was to look at English language proficiency assessment from the world view, life experience and philosophy of the participants. I wanted to know what they are doing with assessment, but also the purposes as they saw them, the conditions in which their practices are conducted and outcomes. I also wanted to know what the concepts, issues and things happening meant to them. I wanted to not only hear their voice, but find out if they had a voice.

Interviewees were prompted and queried in a style that was informal, but not totally lacking in structure or purpose. As Kvale explains, conversation as interview "goes beyond the spontaneous exchange of views as in everyday conversation, and
becomes a careful questioning and listening approach with the purpose of obtaining thoroughly tested knowledge" (1996, p. 6). The interviews were a fine balance between sharing ideas with the participants, carefully allowing the participant to dominate the conversation. He also recommends that “The interviewer’s questions should be brief and simple” (Kvale, 1996, p. 132).

My research design proposed two qualitative interviews despite the fact that the literature primarily supports more than two interviews when using this method. Seidman (1998, p. 11) suggests a three interview model. But for the purposes of my study, two interviews were sufficient for several reasons. The interviews were part of a mixed method study using other methods to collect data. The focus group served as a sort of interview in that it was used to verify information from the interviews. Participants were forthcoming in providing information in the first interview and second interviews. Saturation was reached through the data collected, thus eliminating the need for more data. The same categories and patterns began to surface in the data collected.

First Interviews

Following the initial and return of the Consent Form, an appointment data was scheduled for the first interview. I used a combination of telephone and e-mail to arrange appointments. The location was chosen by the participant. I met most of the participants at their work sites, though.

I took time with this first interview and used field notes along with the recorded interviews to gain information. The field notes were used to help me enter into the world my participants and better understand them. For example, the long drive off the interstate to the small town where one of my participants worked gave me a sense of the isolation
of this rural school district. The constant commotion of children of all ethnic backgrounds coming and going as I waited in the main office for another participant, gave me a sense of the multicultural flavor of the inner city school building in which she worked.

I jotted down many notes, consistent with an ethnographic approach. Perhaps the notes describing the school secretary and gentle, humorous manner would not be significant as I weeded through data and make connections to my research questions, but I was not sure at the time. The notes also served to help remember the day, the interview and some of the nonverbal things that were not captured on the audio tape. “A word or two written at the moment or soon afterwards will jog the memory later in the day and enable the fieldworker to catch significant actions and to construct evocative descriptions of the scene” (Emerson, Fretz, & Shaw, 1995, p. 20). A number of participants made significant statements after I had ended the interview. Several times I captured these comments in field notes and used them to note the issues in the interview data.

I took time to get to know those participants who I had just met or visited casually with those that I knew. I also took a few minutes to review the purpose of my study, answer questions and go over the confidentiality issues. The interviews averaged 35 minutes with the shortest interview being 31 minutes and the longest one being 42. All participants appeared comfortable and were not hesitant to express their ideas and thoughts. The content of the interviews involved general discussion of English language proficiency assessment and the assessment requirements. They also involved discussion of the interviewee’s background and experiences with language minority students and assessment. I summarized data from the questionnaire and we discussed the results.
Seidman explains that "In the first interview, the interviewer’s task is to put the participant’s experience in context by asking him or her to tell as much as possible about him or herself in light of the topic up to the present time" (1998, p. 11). Questions included:

- Tell me about your teaching situation.
- What have you used to assess the language proficiency of the students you work with?
- What works for you?
- What doesn’t work?
- What does English language proficiency mean to you?
- What has your experience been with the new English language proficiency assessment requirements under the No Child Left Behind Act?
- Based on the questionnaire conducted, most of the teachers participating in the survey believed ________ about English language proficiency assessment. What are your thoughts?

Second Interviews

The second interviews were the final data collection activity of my project. They took place in September when all the teachers were back from summer breaks. The interviews were shorter than the first and more focused. They lasted no more than five to ten minutes. Since they were the final data collection, my approach was more purposeful. I wanted to clarify and confirm information gained previously, as opposed to more open ended discussions. The interviews involved specific questions which were based on the previous data. Table 5 provides the questions used in the interviews.
Table 5. Fall Interview Schedule.

<table>
<thead>
<tr>
<th>Name</th>
<th>State</th>
<th>Time</th>
<th>Interview Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amy</td>
<td>I</td>
<td>Wednesday, September 19th at 3:30pm</td>
<td>In your first interview you described all the different assessments that were used in your school district, from teacher developed, classroom assessments to the various standardized tests, including those required for the reading grants you get. You also talked about the challenges of the cultural disconnect with the students and the tests and curriculum and the different values towards education and literacy. What is the relationship between the picture of a student that encompasses his or her culture and community and the picture provided by the large scale standardized tests he or she takes?</td>
</tr>
<tr>
<td>Catherine</td>
<td>I</td>
<td>Wednesday, September 19th at 4:30pm</td>
<td>In your first interview, you stressed the importance of teacher knowledge in determining whether a student has made progress, is proficient enough to be exited and so on. You talked about having a gut feeling about a student, which is based on years of experience and comparison with other students. Could you talk more about what teacher’s know about students? Could you be more specific about what sort of indicators or assessments you would use to find out about a student’s language growth?</td>
</tr>
<tr>
<td>Nona</td>
<td>I</td>
<td>Sunday, September 9th at 3:00pm</td>
<td>You talked about involving students in their own assessments and learning targets. 1. What does assessment that involves a student reveal about him or her that other sorts of assessments don’t reveal? 2. You also talked a lot about formative assessments, multiple measures, portfolios and other things. What do these assessments reveal that standardized tests do not reveal? 3. How could this information be included in the accountability systems that show whether schools are making progress in teaching English and academics? You mentioned the fact that teacher information was anecdotal and not included in the bottom line that the legislators’ and policymakers look at.</td>
</tr>
<tr>
<td>Iva</td>
<td>II</td>
<td>Monday, September 10th at 3:45pm</td>
<td>In your first interview you talked a lot about the importance of observation, running record, home language assessment, self-assessment and other informal and authentic assessments. You said to do a good job, you have to use informal assessments. You also talked about the importance of the required standardized tests and the importance of English Language Learners being part of the accountability picture and not being left behind. 1. Could you explain a little bit more what informal assessment reveals about a student that a standardized test does not? 2. Could you talk about the relationship between these informal, teacher made assessments and the large scale required assessments and accountability?</td>
</tr>
</tbody>
</table>
Table 5. (cont.)

<table>
<thead>
<tr>
<th>Name</th>
<th>State</th>
<th>Time</th>
<th>Interview Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharia</td>
<td>II</td>
<td>Monday, September 10th in the morning</td>
<td>There were several themes in your first interview that I want to pick up on; the issue of what we should expect kids to know and the role and voice of the community in the assessment mandates. You described some of the things that the community would expect of students; knowledge of first language and culture, along with English and being able to Get what you need to get”. You also talked about the fact that the voice of the community is absent in the accountability system and large scale assessment. Could you describe how you could see getting that voice into the requirements and what sort of assessments could support the community’s expectations of what student should know and be able to do?</td>
</tr>
<tr>
<td>Viola</td>
<td>II</td>
<td>Wednesday, September 5th at 2:30pm</td>
<td>In your first interview you emphasized the role of relationships and trust in assessing the English language proficiency of students. Could you explain a bit more: 1. what types of assessments support a relationship of trust and, 2. how do relationships and trust fit into the large scale assessment and accountability requirements? Could you elaborate a little more on your statement when you said “even if I’m not agreeing with the measurement, at least my kids are being measured”</td>
</tr>
<tr>
<td>Wanda</td>
<td>II</td>
<td>Friday, September 7th Lunch at 11:30am</td>
<td>In your first interview you stated that the standardized large scales English Language proficiency assessments were helpful in providing concrete information on students’ skills or lack of skills, particularly in the area of academic English. You also stressed the importance of observation, teacher developed assessments, team decisions, and student participation in assessment. 1. Do you see that these other assessments, including the student’s voice are included in the accountability systems that document progress a school is making towards English language proficiency and content standards? If you, how could these other measures be included in the accountability picture?</td>
</tr>
<tr>
<td>Della</td>
<td>III</td>
<td>Wednesday, September 19th at 2:00pm</td>
<td>In our first interview you expressed a couple key concepts; the frustration with the requirements for adequate yearly progress under the academic achievement assessment program because of the difficulty for your students in meeting AYP goals and appreciation of the fact that your school showed significant growth on the English language proficiency test – TEA test. Could you discuss the role of the English language proficiency assessment and assessment information in the accountability system and picture of your school? Does the fact that the school doesn’t make AYP overshadow your success in English language proficiency? You also discussed the importance of literacy and academic English in looking at a student’s progress towards proficiency and exiting and the fact that many people including the students have an “inflated” idea of their English skills because they become successful in conversational English. Is the state English language proficiency assessment helpful information in affirming the need for instruction in academic English?</td>
</tr>
</tbody>
</table>
Table 5. (cont.)

<table>
<thead>
<tr>
<th>Name</th>
<th>State</th>
<th>Time</th>
<th>Interview Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hettie</td>
<td>III</td>
<td>Thursday, September 6th at 3:15pm</td>
<td>You talked about the importance of communication and working with the parents of your students in your first interview. You also talked about the influence of their culture and home language. 1. How does the information you gain from home language assessment and parental communication help you in learning about the achievement of your students and tracking their progress? 2. You mentioned that your school district was not making adequate yearly progress in academic achievement for the LEP sub-group? Do you know how they are doing in English language proficiency? What are indicators to you that your students are being successful?</td>
</tr>
<tr>
<td>Kim</td>
<td>III</td>
<td>Friday, September 7th at 10:15am</td>
<td>In your first interview you emphasized the importance of the student’s home language in assessment. 1. Could you explain a bit more what assessment of a child’s skills in his or her home language reveals that English language assessment doesn’t reveal and how that information can inform our knowledge of the student? And 2. How do you see information on students’ home language and culture fitting into large scale, mandated tests and accountability?</td>
</tr>
</tbody>
</table>

I used telephone and e-mail to schedule the interviews. I sent an e-mail as a reminder, including an idea of what the question would involve. This gave her an opportunity to think about the answer.

A slight change was made in the research design involving the second interviews. Initially the plan was to conduct two face to face interviews. The plan was changed to a telephone interview for the second interview. The first interviews were successful in providing rich, meaningful data. Therefore, the decision was made with my advisor to eliminate the need to travel hundreds of miles for another interview that could be conducted by telephone.

Focus Groups

The focus groups were conducted as an online chat using electronic media. The purpose of the focus group was to provide a group situation for the participants to share
ideas on the topic of the research. I initiated discussion questions, but consistent with online chats, I allowed the dialogue to flow freely, provided that the issues related to the topic and research questions are being discussed.

The focus groups allowed group interaction for participants in the study and validated data that emerged from the other data collection procedures. I found that the themes and patterns that had emerged in the other data appeared in the focus groups. The group dynamics did not hinder participants from expressing the same ideas. My questions in the discussion were more pointed in an effort to confirm these themes and reach saturation with a different medium.

Initially, this data collection method was planned to include all participants in the same group. In fact several focus groups were held. This allowed for greater discussion amongst the participants and ease in scheduling. Participants were able to pick a time that fit into their schedule. I sent the participants information in advance by e-mail, followed through with telephone calls and also posted information on the Blackboard site.

I used the data from the questionnaires to generate discussion. I also used Robert Tierney’s Literacy Assessment Principles (Tierney, 2005) as an instrument in this activity. This instrument, which can be found in Appendix B, outlines 13 principles or belief statements about literacy assessment developed by Tierney. While the focus of the principles is the literacy assessment of all students, they can be applied to the assessment of second language acquisition. Since they reflect a socio-constructive theoretical base, my interest was to generate discussion from the participants on assessment from this
theoretical framework. Appendix C also provides an example of the correspondence sent out to participants about the focus group.

I used several methods to initiate discussion in the focus groups. As participants entered the chat room I welcomed them and introduced them to the others. I had entered the chat earlier to type out a review of the purpose of the focus group. I referenced this introduction as each one joined. I also referenced the materials that I had sent by e-mail and also posted online. I started with a general question. If the participants continued to discuss the issues, I backed off and simply typed a "great idea!" or some other affirming comment. If the discussion lagged, I posted more questions.

The data collection activities were all completed by the end of September of 2007. Table 4 provided earlier, documents the completion of the final interviews for each participant.

Data Analysis Procedures

My research data consisted of information from all four collection activities as demonstrated in Figure 3. The majority of items on the questionnaire could be analyzed quantitatively. The interview data involved audio tapes and field notes. The audio tapes were transcribed. The focus group data involve transcriptions that were readily available after the chat. These data were analyzed through qualitative coding methods.

The data analysis schedule changed slightly from the original plan in that the collection and analysis activities were ongoing and overlapping. The project originally proposed four distinct data collection activities with a specific order to the collection schedule. Each activity would be analyzed upon completion of each collection. I followed the order established in the research design, but did not wait until I had fully
completed an activity before starting the next. For example, I started interviewing
participants before all the questionnaires had been returned. I conducted focus groups
while I was still completing interviews. I did not complete my final focus group until I
had completed several of the second interviews.

*Questionnaire Data*

The questionnaire, as the first data collection activity, was the first to be analyzed.
The responses from the questionnaires were counted and summarized with the open-
ended questions and additional comments providing information and a backdrop to the
numerical data. I entered the data into an Excel spreadsheet which could then be used to
generate tables. These tables are available in Appendix D.

The questionnaire data, once analyzed, informed my interviews. I could see
where there was agreement or disagreement among my participants. For example, nine
out of eleven participants disagreed or strongly disagreed with the Statement 1
“Information I use to understand my students’ level of English language proficiency is
primarily gathered through formal, large-scale, standardized/standards-based
assessments.” All of the participants agreed or strongly agreed with statement 5.
“Language proficiency assessment information should be used to modify instruction for
students.” I could also see where there were gaps in knowledge. Five of the participants
were either neutral or not very knowledgeable about their states’ English language
proficiency assessment program which was addressed in question 27.

*Interview and Focus Group Data*

A combination of hand coding and computer-assisted coding was used for the
interviews and focus groups. I chose Ethnograph™, which is a computer program
designed to analyze qualitative data, to assist me. I typed out the interviews using Microsoft™ Word processor. Blackboard generated transcripts from the focus groups. Notes and codes were applied by hand to the printed copy of the transcript. The transcripts were then imported into Ethnograph™ for coding and analysis. This program was helpful. I found that the use of a computer program saved me some of the "drudgery of analysis and thereby enable concentrations on meaningful and creative interpretations of what was said in the interviews" (Kvale, 1996, p. 74).

Ethnograph™ also helped me organize the data and my ideas which helped as the amount of my data increased. I was analyzing data as I was continuing with the next data collection activity. "A further advantage is that the programs force the researcher to make explicit commands to the computer, which when reported can give the readers insight into what often seems like a black-box method of interview analysis" (Kvale, 1996, p. 74).

Despite the overlap of research activities and analysis onto the next research activity, my goal of developing theory and meaning as the research progressed was supported. I reviewed my data informally as I moved to the next step. Time was taken between collection activities and data analysis to review theory and literature, revisit research questions and research methods. Each data collection activity informed the next.

Specific examples of where analysis of previous data was used to determine the direction of collections includes the online focus groups and the final interviews. A draft summary of the questionnaire results was posted on the Blackboard website to be used as part of the focus group discussion. I had referred to the results of the questionnaires verbally in my first interviews. It was also available as a document to the participants.
when they logged onto the website for the online chat, allowing the information to be used in the discussions. Final interviews involved specific questions based on first interviews.

Other than the travel to interview my participants, the transcription of interviews was the most time consuming activity. I chose to do my own transcriptions for a number of reasons. I had transcribing equipment that allowed me to use a foot pedal to listen to the digitally recorded interviews that were available as files on my computer. The process was very efficient. There was also the benefit of hearing the interviews again. One of the disadvantages of working on a research project part time is the time lag between events. Memory may fade and field notes are not always sufficient to recreate the experience. Listening to the voices as I transcribed allowed me to remember the setting and re-experience the interview.

Kvale cautions that “Transcripts are decontextualized conversations. If one accepts as a main premise of interpretation that meaning depends on context, then transcripts in isolation make an impoverished basis for interpretation” (1996, p. 167). Kvale (1996) recommends that the researcher do the transcribing in order to secure details relevant to analysis. At a minimum, he recommends transcribing or re-typing parts of the interview which will be subject to analysis.

**Coding**

I used a combination of methods for coding and analysis of interview and focus group data. Huberman and Miles (1994, p. 56) describe codes as “tags or labels for assigning units of meaning to the descriptive or inferential information compiled during a study.” They discuss the use of “descriptive” codes as a sort of first step in providing
meaning to chunks of text. As the investigator goes deeper into the process of analysis, he or she moves into more interpretative and pattern codes (Huberman & Miles, 1994). This concept parallels others who discuss the use of codes, sub-codes, themes, categories and interpretations. Glesne states “Coding is a progressive process of sorting and defining and defining and sorting those scraps of collected data (i.e., observation notes, interview transcripts, memos, documents and notes from relevant literature) that are applicable to your research purpose” (1999, p. 135). I used the following steps in my coding.

1. Initial reading of transcripts
2. Superficial hand coding of transcripts
3. Computer coding using Ethnograph™
4. Review of transcripts for verification of patterns and categories
5. Development of parent codes, categories and definitions in Ethnograph™.
6. Final review of transcripts, Ethnograph™ and analysis of categories.

Samples of Coded Data

I “hand-coded” the interviews during transcription, checking with field notes to make comments and assign codes. These codes were then entered electronically as the transcript was imported into Ethnograph™. An example of a coding of a transcript in Ethnograph™ is provided in Figure 4. Figure 4 shows a screenshot from a piece of text involving several codes. These codes include “ESLLICENSE”, “INNERCITY”, “OVERSEAS”, “PEACECORPS” and “CURRENT”. These codes are a part of the interview from Kim, but they were also used with other participants. Codes are stored in a “Code Book”.

134
Fi

Figure 5 is a segment of my code book. Several codes are seen in this figure, with their definitions. Dates when codes are added or modified are also included. For example the acronym “AYP” was included as a code, with the definition “adequate yearly progress”. It was added on October 9th, 2007. 1STLANGUAGESUP (First Language Support) was added in August, but modified in October. It was assigned a “parent code” at that time.

Ethnograph™ allows for single codes, parent codes, family trees and memos. For example, all of the codes found in the transcript from Kim’s interview were used in some way with other participants. Sometimes the codes became part of another code. ESSLICENSE (ESL Licensure) became part of an education code and a teacher licensure code. The codes, parent codes and family trees provide an opportunity to interpret themes into categories, patterns and themes. Figure 6 is an example of family trees.
Figure 5. Example of Code Book.

Figure 6. Example of Family Tree.
parent code "RELATIONSH" (Relationship) which represents "relationships" has several child codes under it that connect with relationships.

All codes of all transcripts are in the code book under one project. The codes for an individual transcribed interview can also be listed. As more codes are added or more data is coded, commonalities may emerge. The code books and other features of Ethnograph™ help in the organization of data. The parent codes and family trees allow one to see connections. Memos, connected to the project, also provide a vehicle to organize thoughts and notes on the research.

The use of Ethnograph™ did not completely eliminate the need for paper documents with hand-coding and notes. It did assist with organization and provided an overall verification of the trustworthiness of my process of analysis.

Role of the Researcher

The role I stepped into as a researcher for this study was consistent with a naturalistic, grounded theory study. Though my study was not an ethnography, or an anthropological field study, I used methods consistent with ethnographic field research in that I was "committed to going out and getting close to the activities and everyday experiences" (Emerson, Fretz, & Shaw, 1995, p. 1) of my participants. I used other media, such as e-mail, telephone, regular mail and online chats to extend the boundaries between the field and the researcher. Because of the time period of the research and multiple steps, I kept in communication with the participants primarily through e-mail. I attempted to foster a relationship that would support a "professional conversation" (Kvale, 1996, p. 5). The interviews involved an agenda, with specific questions, but I wanted the participants to feel free to discuss their ideas without feeling
they were being quizzed. Though the focus was on their stories and ideas, I did not always refrain from expressing my ideas with the intention of developing a relaxed atmosphere of collegiality and rapport. Occasionally I shared my thoughts and viewpoints. Seidman discusses the balancing act in developing the appropriate amount of rapport that is part of qualitative interviewing. “For the sake of establishing rapport, for example, interviewers sometime share their own experience when they think it is relevant to the participant’s” (1998, p. 81).

Consistent with qualitative research, I made an attempt to get to know the world of my participants as much as possible. Several of the participants had been aware of my research plans since I had first contacted them earlier in my graduate program in preparation of the possibility of needing research study participants. They had some understanding of the project prior to their “official” role as a participant. I spent time on the telephone or through e-mail with the others explaining the project.

All my first interviews were face to face. Most of them were in the participant’s work environment. I spent in the participant’s environment when the opportunity arose. For example, I spent the day with Sharla, visiting schools on her reservation and meeting individuals that she wanted me to meet. There were some visiting professors at the tribal college who were working with the Native language that she introduced me to. The conversations in the car with Sharla allowed me to get to know this participant better and her world view. Even though the casual conversations in the car and during lunch were not recorded as part of the interview, information was reflected in field notes and in my ability to code, and interpret her responses at a later point.
Along with my initial task of explaining the project to the participants, I also explained my role to them. I was clear to state that I was acting in the capacity of a graduate student and was not representing an institution, in the data collection, other than fulfilling my requirements for graduation as a student at the University of North Dakota.

Clarifying my role was most important with those who I had previously known. These individuals were familiar with my employment with the state education agency. In this position, I administer the program for ELLs in the state. I oversee both the state ELL program requirements and the federal Title III program and requirements. Though my actual position is a low level classification within the agency, individuals in the field often interpret my role as more powerful than it is in reality. I wanted the participants to understand that this research was outside that role. Their comments and discussion would not have an impact on their school districts or jobs.

Methods for Verification and Trustworthiness

Validation of data for my research study was provided through a number of methods. Validation in terms of the naturalistic research design I used is defined as assurance of "the correctness or credibility of a description, conclusion, explanation, interpretation, or other sort of account" (Maxwell, 1996, p. 87).

My data was validated in a number of ways. The research design met criteria, according to the literature of appropriate methods and procedures. The use of the questionnaire was consistent with its purpose. My interviews used both uniformity in questioning and discussion, and informality and flexibility, allowing for information to emerge. The focus groups also served as a verification of information gathered through the questionnaire and interviews.
The use of a variety of data collection techniques with the same group of participants served to provide evidence of validity. As a theme emerges from oral interviews, written questionnaires, and group discussion, the theme could be considered valid. Field notes and observations conducted can also serve to document the validity of the themes that I developed from the data.

The steps I used in analyzing the data, including the spreadsheet used in the questionnaire data and the coding procedures I used in the other data ensured validity. The steps followed appropriate methodology as addressed in the literature. The use of a computer program assisted in organization and consistency.

I provided opportunities for verification through member checks. These checks involved the use of findings in each data collection step. I used the questionnaire data to inform the interviews and focus groups. I used the data from other interviews and focus groups in interviews. I primarily used the technique to reaffirm a statement a participant made, such as "other teachers felt that same way as you do." Often this served not only to validate the data but to encourage the participant to go deeper into the issue.

In discussing the methods used in the research design for verification and validity, it is important to discuss the possible threats to validity in this particular study. In reviewing the overall purpose to the study, which is to look at educator perceptions and practices in the area of English language proficiency, oppositions to the development of accurate conclusions in this area would include:

- The participants are not sufficiently knowledgeable in language acquisition or assessment to provide meaningful data,
• The researcher's meaning or understanding of the situation is over-imposed, without allowing the voice of the informant to be heard,

• The researcher's understanding of existing theory biases the data collection techniques and analysis to such an extent that research is compromised.

These threats to validity were addressed through data collection techniques that include triangulation, feedback, member checks, and comparison with other research studies on English language proficiency assessment. These methods do not necessarily guarantee validity, but will assist in ruling out the threats and increasing credibility and integrity of conclusions. Inappropriate conclusions because of the above threats were avoided in the following ways:

• I selected teachers who had an appropriate educational background and experience.

• The possibility of the researcher's own ideas emerging, rather than the research participants' meaning was addressed through the use of appropriate research techniques

• Member checks, in which conclusions and data are reviewed by individuals who had participated in the study were used.

Limitations

The limitations in the study included those factors that had the potential of impacting the study, but were out of my control. The limitations to the study included:

• Challenges of the online chat focus groups:
  
  o Scheduling was difficult,
  
  o Participants had trouble navigating the technology, and
Discussion was not fluid and interactive.

- Participants had busy schedules
  - It was difficult to schedule activities, and
  - Summer vacation impacted the time schedule of data collection.

Despite the limitations of both the online chats and busy schedules, the overall design of the research study was followed and the data was valid and reliable. The essential data collection activities which included the interviews were strong and successful. All participants were thoughtful and capable and provided wonderful, meaningful data to the project. The online chats, though impacted by some issues, were useful in that the data served to confirm themes and patterns that had surfaced previously.

Summary

In summary, the research design as demonstrated in Figure 3, based on a constructivist model of interactive data collection activities, was successful in providing rich data that could be analyzed. All data collection methods served their purpose of providing the sort of data that could add meaning to the next step and the overall project. The mixed method design allowed for triangulation and verification of data. The participants were well-chosen in that they were diverse enough to provide different perspectives and meanings to common ideas and topics. Yet, there were commonalities. The number of participants, variety of methods and other factors allowed for a saturation of data and the possibility of developing patterns, categories and themes that could be interpreted. Overall the research design and methods were successful in supporting a naturalistic research project using primarily qualitative grounded theory methods.
CHAPTER IV

FINDINGS

Chapter IV reports the findings of my research. The chapter includes an analysis of data I collected. As stated in Chapter I, my study examines the problem related to the distance between large scale English language proficiency educational accountability systems and the classroom. The research questions address the assessments teachers are using with ELLs and what these assessments reveal about the students. I discussed with the teachers their viewpoints with the large scale assessments and accountability systems currently in place in the participating states. I explored understandings of proficiency in English with the teachers.

Specifically, the research questions include the following:

1. What assessments are English as a Second Language (ESL), English Language Learner (ELL), and bilingual teachers using to document the development of English language proficiency of English Language Learners?

2. What do the teachers say the assessments reveal about the students?

3. What is the relationship between teacher developed assessment information and large-scale assessment data and accountability?

The research findings are reported in two sections. The first section provides a review of the data that resulted from each specific activity. My research design involves a mixed method approach with four distinct, but related, data collection activities. The
purpose of reporting the data separately is to show how each activity contributed a unique feature to the overall design. Though the data from each activity distinctively reflects the method used, all the data from all four collection activities were used to develop the themes. The themes that were developed from all the collection activities are reported in the second section. Examples of data are used to support the themes. Chapter IV concludes with a summary of the themes and findings.

Section 1: Overview of Data From Each Activity

Each specific data collection activity resulted in data that were unique to the type of activity. As explained in the discussion of the methods, the choice of specific tools and activity was made in order to gather data on the same issues from different methods. Figure 3 in Chapter III clarifies the importance of the different data collection activities, their order in the schedule, and the fact that each step informed the next.

Though each data collection activity generated unique data, distinctive of the method used, the same research problem and questions were addressed. This allowed patterns and themes to emerge as the data were collected. Data were also reviewed against research questions, theoretical basis of study and literature.

The research questions were used as guides in the data collection process. More data was generated that applied to two of the questions. This discrepancy between one of the questions and the data collected will be discussed later. Overall, though, I addressed the questions in all of the data collection activities. Table 6 represents the audit that was conducted as each activity was completed to ensure that the questions were addressed. All questions were sufficiently addressed through the data collection methods as seen in this table.
Table 6. Data Accounting Table.

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Questionnaire (Snail Mail)</th>
<th>Data Collection Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Online</td>
</tr>
<tr>
<td></td>
<td></td>
<td>First Focus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group Interview</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Final Interview</td>
</tr>
<tr>
<td>1. What assessments are ESL and Bilingual Teachers using to document the development of English language proficiency of English Language Learners?</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2. What do the teachers say the assessments reveal about the students?</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>3. What is the relationship between teacher developed assessment information and large scale assessment and accountability?</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

**Questionnaire Data**

As discussed, the data from the questionnaire provided informal, baseline information. I used the information to initiate discussion in the interviews. It is the only data collection activity that uses all eleven since the participant who dropped out after this activity was not replaced.

The questionnaire will be briefly summarized by the categories used in it. It will also be referenced in the discussion on themes. Examples of the data that supports the findings and themes are provided in the text. A complete summary of the questionnaire can be found in Appendix B.

The results of the questionnaire documented several basic ideas about the research participants. These ideas include the fact that the participants:

- Are using a variety of English language proficiency assessment information,
• Have various ideas on how to document that a student has become proficient in English, and
• Are mixed in their experiences and viewpoints on the federal assessment requirements.

Part I of the Questionnaire

Part One of the questionnaire addressed general information on English Language Proficiency Assessment practices. This information provided data for the first research question on assessments teachers are using. The questions addressed the sources teachers used to understand students, how they used that information and what type of assessment information they used.

The responses in the first section reflected a general practice of using a variety of assessments to understand students’ growth in language proficiency, with an emphasis on authentic or classroom based assessments. There was a high rate of agreement on the use of language proficiency assessment information. All the participants either agreed or strongly agreed that assessment information should be used to modify instruction and develop instructional plans as reflected in Tables 7 and 8.

Table 7. Use of English Language Proficiency Assessment: Modification. - Language proficiency assessment information should be used to modify instruction for students (n = 11).

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>7</td>
</tr>
</tbody>
</table>
Table 8. Use of English Language Proficiency Assessment: Student Plans. Language proficiency assessment information should be used to develop instructional student plans (n = 11).

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

Part II of the Questionnaire

Part Two of the questionnaire moved beyond assessment practices into information on indicators of language proficiency. This information supported themes that related to the second and third research questions, looking at what assessments reveal about students and teachers' voices. Responses varied in this area with the greatest divergence being the areas of the role of home languages in English language proficiency and grades. Table 9 documents that the participants varied widely in responding that non-use of the home language is an indicator of language proficiency. Table 10 documents the difference of opinion on the role of grading in determining whether a student is gaining English language proficiency.

Table 9. Student Scores. Student scores “advanced” or “proficient” on a norm-referenced English language proficiency test (n = 11).

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 10. Student Grades. - Student maintains a “C” or better in academic content areas in the classroom (n = 11).

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Part III of the Questionnaire

The results from Part III of the questionnaire, which specifically addressed the NCLB requirements, also documented the variety of experiences and background of the participants. Not everyone reported being knowledgeable on NCLB requirements for ELLs. Though some participants reported having some involvement in their state’s development process, others had not been involved. Table 11 shows that not all the participants had been involved with their state assessment systems.

Table 11. State Assessment Development. - The No Child Left Behind Legislation requires that states involve teachers, administrators and parents in the development process of State assessments. Have you been involved with this development process? (n = 11).

<table>
<thead>
<tr>
<th>Not at all Involved</th>
<th>Not Very Involved</th>
<th>Neutral</th>
<th>Somewhat Involved</th>
<th>Very Involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Please Explain
- I have not participated in item analysis or setting cut scores, but am the district test coordinator.
- Served on western state consortium test development committee; trained to administer WIDA ACCESS; served on alignment committee; served on advisory board
The information reported on the NCLB requirements in academic content tests generated a variety of responses, both positive and negative. The question that addressed the AYP requirements of NCLB resulted in mixed responses. Table 12 shows the divergence of opinion on the inclusion of ELLs in the academic achievement accountability system.

Table 12. Desegregation of Student Scores. – The No Child Left Behind Legislation requires that states and school districts disaggregate the data for academic achievement on students who are limited in English language proficiency achievement goals. Is it useful for you to see how the limited English proficient students do in comparison to other students? (n = 11).

<table>
<thead>
<tr>
<th></th>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Please Explain
- In writing effective reports and setting school goals.
- Yes, I can forget what “average” kids are capable of.

The issue of equity and representation, which would become a theme in the later activities, surfaced in this section of the questionnaire. When asked their opinion on the public reporting of English language proficiency test results, a participant responded that the information is helpful to her because “now our kids count and can’t be ignored”.

*First Interview Data*

The data from the interviews provided detail to many of the concepts that were in the questionnaire. This data includes the description, stories and discussion that the questionnaire data lacked. It includes descriptions of the participants’ classrooms, offices and communities that were a part of my field notes. I also made notes on discussions.
outside the interviews. I gained information from the interviews in areas that were not addressed in the questionnaire.

The participants described the assessments tools they used, as well as their viewpoints and attitudes. They discussed how they used different assessments and the information they gained from these assessments. They talked about their students and explained more about what proficiency in English meant to them and what it did not. Their discussion focused on the large scale and required tests, the federal and state regulations and how it affected them and their students.

The participants had strong opinions on assessment and teaching ELLs. Much of this information was shared in the context of their work, their world and their experiences. All participants discussed their jobs and experiences and the challenges related to these areas. The challenges, obstacles and accomplishments the participants experienced became more of common topic in the interviews than I expected.

The discussions related to the demands of their jobs diverged a bit from the research problem and questions. But, in the interest of developing a rapport and having a conversation, I carried on with the discussion. I found that this became a pattern that was echoed in each interview with each participant.

I continued to explore the topic related to the challenges participants face in their work, carefully trying to make connections with to the world of assessment. The description of their demanding jobs was often confirmed by the notes I had taken. I observed Catherine’s small, closet-sized teaching space when she discussed her heavy caseload. More than one interview with a teacher was interrupted by questions from students or other teachers. I was able to use much of this data in the various themes;
particularly the themes that address teachers' ability to articulate what they know and their voice in decision-making.

Another area of divergence was the concept of culture and home language. Several of the participants, because of their backgrounds, and current roles as heritage language teachers, discussed the importance of the student's home language in his or her educational success, including English language proficiency. Because this concept emerged, I addressed it with others and continued to explore it.

Finally, as I continued to explore the concepts of English language proficiency assessment, assessment requirements and teachers' involvement and viewpoint, I realized that some of my participants were able to articulate information more clearly than others. I reviewed this finding, which surfaced in the interviews with my advisor and we discussed it in terms of social constructive theory which relates thought and language. I continued to explore this concept related to a teachers' articulation of assessment knowledge in the rest of the data collection activities.

Overall, the first interviews, which were, as discussed, the heart and nucleus of my project. They validated and provided detail to the themes that had begun in the questionnaire. They provided information that could be further explored in the focus groups. They also served to provide data in areas related to the research problem, but were not specifically addressed in my questions. These areas were also further explored in the focus groups and final interviews. The data assisted in providing greater meaning and significance to concepts and patterns that were developing. My four themes were drafted from these interviews and further developed with the completion of the subsequent activities.
**Focus Group Data**

The focus groups, consistent with their purpose, provided data that allowed concepts to be developed and confirmed in a group situation. I found that the participants expressed the same concepts that they had in the interviews. Though they were discussing issues online with individuals they had never met before, there was a lot of common agreement. The statement “I agree, Hettie” “I agree, Wanda,” and “Sharla, you are so right!” were helpful in providing confirmation to me of patterns and themes that had emerged.

The focus group activity presented the greatest challenges of the data collection activities because of scheduling and technology issues. All participants have very busy lives and coordinating schedules was difficult. Despite advance instructions, a number of individuals were not able to access the Blackboard website where the chat was being conducted. They contacted the technical assistance staff at the Center for Instructional Technology and were able to get online, but time was lost. I worried that this loss of time and the stress of having problems would impact the data negatively.

However, once analyzed I found meaningful data. The voices of the participants are strong and expressive as they supported each other and, in some cases, expanded on an opinion or idea that another had expressed. There was evidence of positive group interaction. An example of this is in the following dialogue:

Viola: I think teachers get their understanding of progress mostly from the work they do in the classroom. Large scale is less valuable – results take too long to get.

Della: I agree.

Mari: Both of you seem to see a disconnect between the teachers and the large scale tests.
Della: Absolutely! They are much too remote by the time the results arrive.

Viola: Definitely. We are trying to use the data more in our district and are making progress, but it is still the time thing. Also the time involved in analysis....

The participants were most expressive on the importance of their own authentic assessments in reporting student growth and progress. I asked specific questions about their use of assessment and how they perceived the relationship between their view of students and assessments they used as opposed to external assessments.

They were supportive of many of the principles stated in Tierney’s Principles of Assessment. As discussed in Chapter III, these Principles were sent to the participants by e-mail and also posted on the Blackboard website. They can be found in Appendix B. As part of the focus group, I asked the participants to respond to the Principles in view of their own assessment practices and beliefs. I asked them what they thought of the Principles. I asked how they compared with their own beliefs and whether they could be applied to multicultural and ELL students. It did not take much prompting for discussion to develop in this area.

Some of the comments related to these principles include the following:

“I can tell you that I love Tierney’s Principles and they are compatible with my beliefs,” Iva states.

Hettie agreed with her.

“Yes, I am impressed with those principles – especially ‘assessments should not be imposed on the classroom.’ This is what is frustrating.”

The concern expressed over the federal requirements and large scale assessments being the only voice in accountability was also strongly expressed in the focus groups.
"The policy makers need to listen to teachers, the people in the trenches - we taught them," Sharia explains to support the need for teacher involvement.

Overall, the focus groups served to validate the findings that had previously emerged. They were clear on some of the issues that had been briefly addressed previously, such as the need to include the teachers’ voices in accountability decisions. They affirmed and supported each other, allowing for validation and confirmation of the patterns and themes.

Second Interview Data

The data that emerged from the second and final interviews, similar to the focus groups, served to validate and confirm. Since all the data from all activities had been transcribed, coded, analyzed and categorized, I used the final interviews to check on some of my findings. Table 5 provides a summary of my questions.

For example, I was struggling with the data that had emerged concerning the teachers’ explanations on how they know students are progressing in English and academics. “Gut” knowledge was referenced and while some participants were able to more clearly articulate what that gut knowledge of student progress and proficiency was, others were not. Since this issue of articulation was an emerging theme, I explored it more specifically with some of the participants.

Catherine provided more information for me on what her gut feeling on student progress means by describing her years of experience along with some of the tests she uses and how she interprets them. She refers to a test her district uses for literacy assessment. “So, teachers use Degrees of Reading Power (DRP) to look at reading, but it also shows us if there’s a kid we’re wondering about and is LEP - that certainly would
give us ideas." She told more stories about her students. She summarized the interview by explaining more about what goes into her process of evaluation student growth. She explains that the gut feelings are backed up with information from testing, such as the criterion referenced test (CRT) that the state mandates. She also explains that the fewer accommodations a student needs on a standardized test are a clue that he or she is becoming more proficient in English and in less need of support through the ESL program.

Yeah, I mean certainly every teacher has gut feelings, but we have to back it up with information. We provide students with assistance when they take the CRT or they take any sort of national test. There are certain accommodations that you can make for LEP kids. Like I said, if they need fewer accommodations, then you know if they have reached a point where they’re more independent. The independence is shown by needing less help in the classroom and being more independent in their work and so that’s one indication – but it’s not just ‘gut’ – you’re actually seeing on paper that they’re improving in their assignments and other things.

Overall, the findings from the second interview provided further detail and description. Areas that that were not completely clear were clarified. Observations from the data that I was working with were confirmed. In that regard, the second and final interviews served as a sort of “member check” of some of my findings. I had reached saturation in that I was hearing the same things again. I was able to move forward and complete the development of the themes that had emerged.

Section 2: Themes

I learned a number of things from each data collection activity. As discussed, even though each activity reflected the uniqueness of the approach, they all addressed the same research problem and questions. Therefore, concepts, patterns and categories began to surface as I continued with the research. These ideas include the following:
• Teachers who work with English Language Learners use a variety of assessments to learn about their students' development in English language proficiency.

• The different assessments reveal different aspects of student growth, development, language use and proficiency.

• Teachers approach instruction and assessment from a variety of backgrounds, experiences and understandings.

• Teachers know what their students can do and cannot do.

• Teachers have mixed feelings about the state mandated tests for assessment and accountability.

These statements, which summarize some of the findings from my research study, may seem ordinary and mundane. There are also paradoxes and contradictions in my findings. Teachers know things about their students and yet, the manner in which they articulate that knowledge varies widely. Teachers participate in administration of large scale assessment. But the teachers do not always use the information gained from these assessments in their practice. Teachers use authentic assessment, but this information is not necessarily included in the large picture of how the students are doing. Teachers are both acknowledged by policymakers as key to the success in children's achievement, yet not necessarily given a voice in assessment and accountability decisions. Teachers' voices are also impacted by their schedules, case loads, positions and other factors.

Gaps in Findings

Along with paradoxes and contradictions in my findings, there are also gaps. Some of my research questions are addressed better than others. I did not obtain much
data to provide information on my second research question which relates to what the teachers say the assessments reveal about students. I struggled with this issue through the data collection process since part of my process was to review previously collected data, literature and theory between each step. Part of the review of literature included research texts. In reviewing the literature and reflecting on my procedures, I made several conclusions.

I started with a sound research design with questions that served as guides in which to enter the research field. The literature on naturalistic research designs emphasizes the need to start with a framework and research questions, but to be open to possibilities.

Although early identification of the research question and possible constructs is helpful, it is equally important to recognize that both are tentative in this type of research. No construct is guaranteed a place in resultant theory, not matter how well it is measured. Also, the research question may shift during the research. (Eisenhardt, 2002, p. 11)

Consistent with this approach, I made efforts to gather data from the participants on issues related to the second question, but did not pursue the issue at the expense of losing the quality of a professional conversation. The lack of data could be looked at from a number of perspectives. The first perspective would consider that this lack of data is a finding in itself and relates to the theme of articulation of assessment information, which will be addressed. Another perspective is that the question may not have been suitably phrased for the research project. It may have been more effective if I had included observations or student artifacts.

The final perspective is that whether the question is answered or not answered is irrelevant in a naturalistic study. The questions are simply guides to use to enter the field.
in an attitude of discovery. Whereas a positivistic approach would expect an answer to each research question, a naturalistic study should be judged on the quality of design and methods.

Overview of Themes

The data fall into four themes. Figure 7 portrays the themes in a visual context, using a similar display as used in the research design showing background issues at the bottom. The themes are displayed against a backdrop of topics that have influenced the research. These background topics, presented in a green font, include such things as increasing numbers of ELLs in the schools, increased emphasis on accountability nationally and other issues that impact the education of multicultural students today. The phrases in the blue font represent some of the codes and categories that contributed to the themes.

The themes flow together with the first theme addressing more concrete information about assessments. The themes also follow the research questions in that the questions move gradually from information about students and assessment practices, into information about teachers, policy and relationships.

The first theme, which includes the types of assessments teachers use, includes three sub-themes. These sub-themes address the different assessments teachers use, what the assessments reveal and how teachers understand language proficiency assessment and student development.

The second theme addresses what teachers know about what students can do. The main theme involves information on what teachers know. There are two sub-themes
Theme 1:
Teachers use a variety of assessments to understand English Language Learners

1a. Teachers use formal and informal assessments to document student learning and progress towards proficiency

1b. Teachers claim that both formal and informal assessment can reveal important information about students

1c. Teachers' perspectives on English Language Proficiency and Program Exit Criteria Depends on Contextual Factors

Theme 2:
Teachers know what their students can and can not do

2a. Teachers clearly express that they know their students

2b. Teachers articulate their knowledge of students with varied clarity and often through stories

Theme 3:
Teachers bring a variety of backgrounds and contexts to the assessment process when assessing at English Language Learners

Theme 4:
The relationship between teacher developed assessment Information an large scale assessment and accountability is multifaceted and mixed

Limited funding
Marginalization
Time
Limited funding
Gut feelings
Increased emphasis on accountability
Funding
Marginalization
Increased need for citizens with high levels of English literacy and proficiency

Greater need for citizens with high levels of English literacy and proficiency

Teachers' voices lost in broadcast of federal mandates

Figure 7: Research Themes and Findings
within this theme and they address the more complex issues of how teachers articulate
and express the information they know about students.

Theme 3 continues with the issue of student assessment, but moves more into the
influence of the teachers on the assessment process. It addresses the background and
context that teachers bring to assessing students and interpreting that assessment.

The fourth theme focuses more on the teachers and addresses the environmental
issues, such as the large-scale high-stakes tests and accountability. It includes
information on the relationship teachers have with these assessments, their perceptions,
and viewpoints. This theme is both simple and complex and addresses the center of the
research in that it looks at the voice of teachers in assessment and accountability systems.

The following discussion reviews the themes, relating them to the research
question and data collection methods.

Theme 1: Teachers use a Variety of Assessments to Understand
English Language Learners

The first research question addresses the assessments that ESL and bilingual
teachers are using to document the development of the English language proficiency of
English language learners. Theme 1 provides information for that question. Data from
the questionnaire, interviews and online focus group produced information documenting
that teachers use a variety of assessments, from informal, observation and performance
assessments to formal, standardized and large scale language proficiency tests. This data
developed into a theme which, as stated, appears mundane and insignificant, but has
significance in the details of how teachers reported this information, what the
assessments mean to them and what they reveal about students.
The participants explained that the different assessments they use reveal students in different ways. This difference in how teachers document development and attainment of language proficiency is reflective in the types of assessments they use and value and how they view their students. It also reflects their understanding of what it means to be proficient in English.

**Theme 1a: Teachers use Formal and Informal Assessment to Document Student Learning and Progress Towards Proficiency**

Table 13 lists the various assessments reported by the research participants. The data used to put the table together came from all collection activities. In the interviews, the participants provided the most detail on the individual assessments they used.

Table 13. Assessments Used by Participants.

<table>
<thead>
<tr>
<th>Authentic and Classroom-based Assessments</th>
<th>Standardized, Commercial Assessments</th>
<th>Large Scale, Standards-based Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informal</td>
<td>Formal</td>
<td>Formal</td>
</tr>
<tr>
<td>• Observation</td>
<td>• Assessments related to</td>
<td>• State English language proficiency</td>
</tr>
<tr>
<td></td>
<td>instructional programs, like</td>
<td>tests - all participating states</td>
</tr>
<tr>
<td></td>
<td>Compass Learning</td>
<td>have different programs</td>
</tr>
<tr>
<td></td>
<td>• Dynamic Indicators of Basic</td>
<td>• State academic achievement</td>
</tr>
<tr>
<td></td>
<td>Early Literacy Skills (DIBELS)</td>
<td>assessment tests</td>
</tr>
<tr>
<td></td>
<td>• Fox in the Box</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Gates MacGinitie Reading Test</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• IDEA Proficiency Test</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Language Assessment Scales</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Mainstream reading assessments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>associated with specific commercial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>reading and language arts programs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Measures of Academic Progress</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(MAPS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Woodcock Munoz Language Survey</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Student self assessment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Student Oral Language Observation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Matrix (SOLOM)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Running Record</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Parent information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Portfolio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Games</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Guided reading</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Learning Record</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• CBMs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Checklists</td>
<td></td>
</tr>
</tbody>
</table>

161
The table is divided the two categories, which are then further divided into the categories of authentic and classroom-based assessments in the informal category and standardized, commercial assessments separate from large scale standard based assessments which are both in the formal category. These types of assessments were addressed in Chapter II in more detail.

Many of the types of assessments that the participants use with ELLs that are considered informal assessments are the same assessments that are used in the mainstream classroom. Practices such as student observation portfolios, anecdotal notes and checklists are authentic assessments that originated with mainstream educators.

Some of the authentic assessments used by the participants that are unique to ELLs include the SOLOM and first language assessment. SOLOM is the oral language observational tool that was discussed in Chapter II. First language assessment could include more formal tests, but it was included in the informal category since the participants that mentioned it used it in a more authentic manner. The participants stated that most of these authentic, classroom based assessments were used in a formative manner to document progress and gather information to inform instruction.

The use of informal assessment varies according to the teacher. The data revealed that nearly all participants express an appreciation and belief in this sort of assessment. Nearly all responded that authentic or performance assessment was more helpful to them when asked on the questionnaire about the types of assessments. Table 14 documents the reporting of authentic assessment on the questionnaire.
Table 14. Helpfulness of Informal Assessment. – What assessment information on your students is most helpful for you? Informal assessments that are authentic or performance based (for example - running record, observation, miscue analysis, etc.) (n = 11).

<table>
<thead>
<tr>
<th></th>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Kim from State III explains that one of the difficulties of the state English language proficiency test is that the teachers don’t get the results as soon as she would like. She explains that the testing information is helpful, but “not as helpful to me, though, as what I do daily.” She explains that the daily assessments she uses involve a district assessment they have adopted to assess fluency. The acronym for this assessment is “CBM”. Kim states that she uses “my own assessments or observations, or the words per minute - we do CBMs – words per minute in Spanish - and in English and comprehension tests, and just observational rubrics.”

The participants reported using many standardized, commercial assessments with their students. This category, which is included in the formal assessments on the table, includes many of the traditional norm-referenced tests used to document English language proficiency. The Woodcock Munoz Language Survey, Language Assessment Scales and Idea Proficiency Test are all norm referenced English language proficiency tests published by large testing companies. Most of the teachers in the study continued to use the traditional English language proficiency tests along with the new state developed tests required by the federal legislation.
The other assessments listed on the table in the category of standardized, commercial assessments are tests designed to assess the literacy or academic content skills of mainstream students. These tests were often used because of specific program requirements or district curriculum adoptions. ELL teachers that work closely with mainstream classroom teachers use many of the same assessments used in the mainstream classroom to assess the progress of the English language learner or compare against mainstream student progress.

Amy explains that the school district that she works in has benefited from many federal grants because of the students' low achievement levels. The district has been using a number of assessments required by these grants and training on the implementation of the assessments. She explains that the district uses Dynamic Indicators of Basic Early Literacy Skills (DIBELS) for the Reading First Grant. Reading First provides federal funding under Title I of NCLB for early literacy. It requires a very prescripted, skills-based approach to literacy. DIBELS provides information on letter/sound recognition in English speaking students. It was not necessarily designed for language minority students, but is a common instrument used by districts that receive Reading First Funding. This is one of many assessments used in this district to evaluate students’ language skills and progress, she explains. Other participants use this test too, as reported on the questionnaire.

The large scale standard-based tests listed in the formal category of the table include the tests each state had chosen to meet NCLB requirements in the areas of academic content standards and English language proficiency standards. Each state had different standards and different assessments. These tests were being implemented in all
the states in various stages. The participants had different levels of experiences with the tests. Some teachers reported using the data from these tests more than others.

Information from tests used to assess achievement in the content area is also used by the teachers of ELLs. The language arts portion of the state academic achievement test is often used to help determine proficiency and growth in English. Amy explains that before the state English language proficiency test was implemented the district used the tests that were designed for other purposes to identify students as English language learners. "We have a criteria sheet with four different things on it and then it goes up to the buildings and the buildings put the scores on and then they have a team of people that decides."

The participants reported that the use of the state English language proficiency test results was minimal. Because the tests are still new and two of the states had gone through only one administration of the test at the time of my research, teachers did not have as much familiarity with the tests. Catherine had helped develop the test that her state (State I) was using, but was disappointed in it. "I'm not happy with it. It is not a good test. It's too easy, basically." She doesn't plan to depend on the test very much to learn about her students. She does see hope for improvement. "I am sure we're going to tweak it and fix it and over the years, it will improve. But, this is the first year, what do you expect? We did our best."

Iva, who is from State II, by contrast, expressed excitement about the use of the new state English language proficiency test which can be used along with her informal assessments to complement what she knows. "Now we are adopting, implementing the WIDA ACCESS assessment for English language proficiency which is really good in
comparing.” Wanda, who is also in State II, echoes the advantage this test will bring in providing another piece of information in looking at students. “I think that it will be really neat to see the results and see how that corresponds with the proficiency that these kids had before on the Woodcock Munoz.”

Most of the teachers that participated in the study used assessments that are included in all three categories to learn about their students’ English language proficiency. They explained that choices are influenced by the assessments that are supported by their school districts and programs. They also said that they use assessments that they had learned in their teacher training coursework, workshops and their own experience of what works and what doesn’t work. Ten out of the 11 respondents on the questionnaire either agreed or strongly agreed with the statement that “Information on my students’ level of English language proficiency is primarily gathered through a combination of large scale and standardized tests and authentic or classroom based assessments”. The interviews and online focus groups supported the fact that the participants use a variety of assessments to learn about their students.

Theme 1b: Teachers Claim That Both Formal and Informal Assessments can Reveal Important Information About Students

Theme 1b expands upon the overall topic of the variety of assessments that teachers use to assess ELLs and focuses on what the different assessments reveal about students. The participants reported throughout the study that they use different assessments because of the different purposes they serve and the information they provide. They explained that different assessments reveal different aspects of language
development and ability. A complete picture of a student is revealed in using multiple assessments.

Though all of the teachers participated in the large scale state tests in some manner and also used commercial tests, none of them used these formal assessments for all their information on students. Most of them supplemented the information they received from English language proficiency testing with other sorts of assessments. This use of supplementary information was initially reported in the questionnaire and expanded upon in the interviews. Table 15 documents the use of additional information.

Table 15. Supplementing of Assessments. – How often do you supplement English Language proficiency test scores with other sorts of assessments to understand what students need and how they are doing? (n = 11).

<table>
<thead>
<tr>
<th>Never</th>
<th>Rarely</th>
<th>Seldom</th>
<th>Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

This interest in using both informal and informal assessment was reflected in different ways, based on the context of the teacher. Kim, who is a Spanish literacy teacher, emphasized the home language information, along with English language assessment. Della, who works with high school new immigrant students, emphasized that a standardized language proficiency test score is important information to include with classroom assessments when teachers might be fooled by a student’s ability to converse easily. Though the context may be different, the belief in a need to use both informal and formal assessments formed a pattern that was consistent with nearly all the participants.
Though they valued both, the participants did emphasize the greater importance of informal assessments. They felt that these assessments present a more complete picture of a student and are more compatible with instruction and appropriate for students of different cultural backgrounds.

Iva, an experienced K-12 ELL teacher from State II, explains the overall importance of assessment when she states “I think assessment drives instruction.” She goes on to explain the importance of informal assessment: “I really am a true believer in differentiated instruction and that really takes more than just a standardized assessment and paper and pencil and multiple choice assessments.” She explains, “So, I do a lot of observing, checklists, running records...self-assessments.”

Nona, who is from State I expressed a similar thought as Iva, stating:

I don’t believe in testing, testing, testing... just for the sake of testing. I think what we can do is we can look at other things without having to give more than one test. I really think that we can talk to teachers to see how long that student is maneuvering in a classroom and then observe that student to see how easily they work and interact with the textbook or interact with the students when they’re doing like a lab or project activities.

Nona explains that the informal assessments involving observation, conversations with classroom teachers and performance assessment reveal aspects of a student’s language development that is not revealed by the test scores.

Student self-assessment is a type of authentic assessment that the teachers value because of the different sort of information about students that is revealed. Students know things about themselves that they don’t express on pencil and paper tests. Iva continues to support Nona’s comments and her ideas on self-assessment. She explains
the value in the information that a child will provide about his or her own growth in English:

Because, you know it’s very important for a child, and it doesn’t matter how much English that child knows - they all really know how much they know - for them to self-assess – to say, ‘O. K. These are the things that I’m really good at and these are the things that I have to work on.’

Nona also supports that fact that kids can assess themselves in providing an actual example from her work. She describes a project in which the students were expected to assess themselves. “Most of the time the kids were really honest about assessing themselves in terms of whether they had really good examples – they seemed to know the areas that they were the weakest.”

Nona continues on to explain a project that involved the students in setting rubrics and assessment themselves. She explains how the students were more proactive about improving when the assessment was in their own hands. She says that the kids:

wanted to actually go back and find an example that was stronger than what they had originally given because they realized afterwards that they hadn’t really fulfilled that part of the bargain – which to me was really powerful because usually when you give kids a test, a teacher will go through and mark it or give them a grade and you give it back to them and that’s it.

Home or first language assessment is also important, according to several participants. Assessment of skills in the child’s first language is a type of authentic assessment that reveals information about a student that cannot be found in formal assessments. Kim explains how observing how a student uses Spanish in both conversation and literacy can inform the assessment process of the child’s English proficiency.
If a child is having difficulties in Spanish, the teachers would realize that these same difficulties in English may indicate a deeper problem than just normal stages of second language acquisition, she explains. “But, there may be a student that is having many, many reading difficulties in Spanish which they speak and think in and hear at home as well as in English, and that would be, then, a red flag,” Kim states.

Overall, the research participants appreciate the fact that authentic assessments, such as portfolios tended to include students’ strengths and abilities, along with weaknesses. Formal and standardized tests, whether they are commercial, norm-referenced tests, or state developed, standards-based tests, don’t always give a true picture of a student’s strengths.

Sharla, from State III, explains how the Learning Record which is a portfolio type assessment that had been used in her school included a whole picture of the student and involved the student. She explains that the portfolio showed the students best work. Even though there were students who didn’t work much and didn’t produce much, they still had a folder with good work:

It was showing their best work, and if the student had a thinner folder - it’s usually the student that was doing better and the students that weren’t doing so well had these thin folders - yet it was their best work.

The research participants emphasized the role of culture when discussing different types of assessment and what they reveal about children. Several reported that culture can influence test scores and student performance. Students of different cultures perform differently on assessments because of cultural influences according to some of the participants. Many saw students from minority cultures as less adept in test taking skills;
thus these sorts of assessments were not as favorable to the students. Amy’s comment is representative of the comments of most of the participants. She explains:

So, they can’t take the tests that are written in academic English and really show what they can do. In their culture, too, they’re not as verbal in English. When I go to a teacher meeting, the Crow speakers – they’re talking very fluently in Crow to each other, but then when they revert over to participate in English with the group, they don’t do it as much. They’re quiet, shy, reserved. I don’t know what it is – but they don’t participate as much.

Amy’s comment on the fact that students may have strengths and abilities that are not reflected in the standardized tests corroborates Sharla’s statement that an authentic portfolio assessment reveals more complete information about a student.

A number of the teachers address English language proficiency assessment in terms of affective qualities, such as relationship and trust. Assessment that is conducted in a safe, trusting environment is more apt to reveal more about a student. “when they trust you and trust that they’re safe in that classroom and then they feel free to ask questions and show you that they don’t know things and try their best,” Viola explains.

Even though all the participants emphasized the importance and value of authentic and informal assessment because it can provide a more realistic, well-rounded picture of the students, they also acknowledged that standardized testing had a place in the educational program of ELLs. A norm-referenced commercial test or a state standards-based test could reveal information or provide documentation of a student’s needs in a way that an observation or portfolio or some other form of authentic assessment could not demonstrate as easily.

Wanda provides an example that documented that the standardized test score shows teachers that students are less proficient in English than previously thought.
Wanda explains that mainstream teachers assumed that the student knew more than she did because she was quiet and polite:

I think they thought that she was a lot higher as far as her comprehension goes than what she was. She seemed to understand and she nodded her head, “yes,” and would say “yes, I understand.” Then someone said, “I don’t think she’s really getting it.”

When Wanda administered the commercial English language proficiency test, the results showed that the student was at a low level of English language proficiency. Thus, it was the concrete, “black and white” numerical test score, based on national norms, that provided the necessary documentation to convince the teachers that the student’s English skills were not that strong and she needed more support in the classroom.

Della explains how a standardized English language proficiency test can show teachers and students in the high school for new immigrants what students know and don’t know. She says that the students, who are often older than average and are anxious to move on and get jobs, pick up on conversational English quickly, which can be deceiving. Just as the mainstream teachers that Wanda works with had misinterpreted a student’s polite behavior for proficiency in English, Della’s colleagues are not as knowledgeable about language development. They do not realize that students can pick up on some verbal behaviors and simple English that fool many people in thinking they know more than they do. Della explains about that both the teachers and the students misjudge the students’ actual proficiency levels.

They get very, very orally proficient and they can come in and really ‘wow’ the teachers – especially the non-ESL teachers. These teachers will say “Well, this kid is ready. They don’t need ESL.” And then you know they can’t read and they can’t write and it’s just a real dilemma for the students because then they have an inflated idea too of what they can do.
Della explains that it is difficult for the students to realize that they have to work harder to gain the academic English that is necessary for them to succeed in school and have opportunities after graduation.

It's really hard for them to step back and realize that they really need to have a lot of serious work in reading and writing in order to bring the academic skills up.... so that they can go on to college or do what they want to do.

The standardized language proficiency test provides the information to show that these students may be adept in basic conversational English, but lack academic language skills, according to Della.

In summary, the participants expressed the fact that they learn different things from different assessments. There was value to each assessment because of the unique information it brought to the table when looking at student growth and development in English.

Theme 1c: Teachers’ Perspectives on English Language Proficiency and Program Exit Criteria Depends on Contextual Factors

As discussed in Chapter II, ELLs are identified as lacking in English through identification assessments. They are then provided support until they demonstrate proficiency or attainment in English. States set definitions of proficiency for federal reporting and accountability purposes. Teachers who work with ELLs must also have an understanding of progress and proficiency in order to provide instruction, interpret assessment and make recommendations on students. Students must be eventually “exited” from the status of ELL.

As discussed, the literature on what constitutes proficiency varies. The concept of “academic” English proficiency has become well accepted with the emphasis on the need
for academic language skills needed for success in the classroom. Whereas many have simplified the issues of language development into the conversational English that children first develop and the academic language proficiency skills that come later, others emphasize that this is too simplistic and doesn’t account for the sophisticated language needed in different social situations (Bailey, 2007).

For bilingual learners, English language proficiency would also involve the use of different languages, dialects and registers in different domains. A social constructivist view of language proficiency encompasses the interdependence of language with culture and social relationships (Fayden, 2005). Ultimately, a definition of proficiency would need to include success in all language areas, including lexical, grammar and discourse, along with written and oral language in academic situations and other areas of human interaction.

I found that the documentation and expression of what is means to be a proficient student in English varies among my participants. As the literature documents, this concept, which seems very simple and ordinary, is, in fact, complex. The findings from my research support not only the fact that it is complicated concept, but also that it changes with the context in which the teachers are working. My research participants have their own ideas, educational philosophies and viewpoints. They also work in real situations, fraught with challenges of scheduling, requirements and caseloads. The data support that finding that proficiency in English means different things to different people.

The area of the questionnaire that had the greatest diversity in responses is the section that addressed Indicators of English Language Proficiency. Though this variety could be due the nature of the questions themselves, it is also a sign that there is not a
common understanding on what proficiency involves. The only area that there was
general agreement on the questionnaire in the area of English language proficiency was
the item addressing the use of abstract concepts in English. All participants agreed that
this is a strong indicator of proficiency in English as Table 16 indicates.

Table 16. Abstract Concepts. Student Discusses Abstract Concepts in English (n = 11).

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>9</td>
</tr>
</tbody>
</table>

The areas of greatest discrepancy include those that address conversational
English, home language testing and grading. Table 17 indicates the lack of consensus on
the role the home language plays in English language proficiency.

Table 17. Home Language Usage. Student quits using his or her home language (n = 11).

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 18 shows that the participants did not agree on the role of state tests in
determining proficiency in English.
Table 18. Proficiency Scoring. Student scores “proficient” in Reading/Language Arts on the state standard-based academic achievement assessment (n = 11).

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 10, which was discussed earlier, also shows that the participants did not agree on the role of grades in determining proficiency in English.

When asked about proficiency and attainment of English in the interviews, I received different responses. A common issue addressed by the participants was the importance of becoming proficient in the type of English needed in the classroom. This “classroom English” is the cognitive academic language proficiency that is addressed in the literature and often referred to as CALP. CALP is the acronym that Cummins (1984) used for cognitive academic language proficiency. Literacy skills, which are also important skills for success in school, along with the ability to discuss abstract concepts in English were reported on the questionnaire as strong indicators of proficiency.

Participants differentiate between the social, conversational English that children learn quickly and the academic English that takes longer. Kim discusses the differences between the two types of English and the importance of both.

So, I think it’s important to be aware of those two different types – and aware that just because a child seems very fluent on the playground or in oral conversation does not mean that their thinking and reading and writing and listening skills are at the same level. So, I would look at how they interact with other children and with teachers in English and all four of those modalities; reading, writing, speaking and listening.
Other than the commonality of emphasizing academic English and the need to be successful in the classroom, participants emphasized different aspects of language proficiency or different ways of exiting students. Some were more articulate than others. Some worked in situations where the district had formalized a set of criteria for exiting. Others did not have formal criteria. On the whole, most of the participants appeared more involved with the instruction of the students at the moment and did focus much attention on concepts related to proficiency in English.

Iva, whose district was enrolling more students on a very limited ELL program budget, emphasized her busy schedule and even though she wanted to spend more time focusing on students who were becoming proficient, she didn’t have time. Catherine also emphasized her busy schedule and the fact that students were exited when her case load became too full. She also did a lot of observing of her students and emphasized the role of experiential knowledge on the part of the teacher. Hettie, another isolated ESL teacher in a small district, emphasized her busy case load. The ESL teachers who worked as resource teachers did not always articulate a formal exit policy, but based their services on needs and time.

Comparison with mainstream students was another indicator of language proficiency. Catherine, who works closely with mainstream teachers explains that “doing well, performing at the level of the peers, the regular American peers in the classroom… average, above average work,” are indicators for her that the student is proficient in English. Nona also observed her students and relied on information from the student and the parents.
A number of participants stressed the need to use multiple measures to document English proficiency and exit students. Wanda states:

I think there are so many different forms of assessment out there and I think if we’re only looking at standardized test results, like with these kids, we are making a mistake, because there are so many other assessments …. and self-assessment… especially for exiting a student from the program.

Viola emphasized that students need to work at grade level, meet achievement expectations and also meet proficiency levels on an English language proficiency test.

Nona explains how authentic assessment, as interpreted by a teacher who knows the student is important in determining whether a student is proficient or not. She describes two of her students and how it was important to observe them to see how they were maneuvering in the classroom and interacting. “I think you can work your way around situations if you are proficient – even if you don’t have the vocabulary. You can’t get that with a test. You have to watch them.”

Along with the variety in defining full proficiency in English, the participants, they struggled with the articulation of a formal set of criteria for exiting students from the status of being English Language Learners. This challenge seems to differ according to their different situations.

Della explains that proficiency in English not an issue that she needs to be concerned with in her school because the high school students are newcomers and most do not have the opportunity to become proficient in English before they leave. “They are probably going to age out before they exit,” she states.

Catherine, who is one of two ESL teachers in a district with a growing population of ELLs, also struggles with her exit criteria because it is based more on available
resources than student need. She “exits” students when she no longer has time to provide services because of her caseload. “Now, in another state they might keep him on longer, because they have the personnel for it. We don’t. I can’t be with everybody. So I have to prioritize,” Catherine explains.

Work load was often mentioned when I asked about criteria for exiting students. The participants expressed that they were too busy teaching and all the other duties they were responsible for to think about such things. School districts did not always support them in looking at assessments other than the large scale required tests.

Kim, whose room was alive with color and decorations and educational materials in Spanish, admits to being “not totally knowledgeable about some of the testing we do.” She also sees a weakness in the system in not having a consistent assessment program related to instruction. She explains that they don’t have a “curriculum with assessments that go with it so that it’s not just one big state test, or district test at the beginning and the end of the year, but rather assessment that’s based on what the child produces daily, weekly, and monthly and you can see growth in that way”.

Amy, whose role was curriculum and assessment coordinator, was one of the few that could articulate the multiple criteria they used in her district to identify students and exit them. The district uses a number of assessments, primarily from mainstream educational programs, along with the state English language proficiency test, to exit students.

This variation in perspective on proficiency in English related to the participants’ positions and the population of students they services. As Della, who worked with older than average students, had explained, getting the students to proficiency wasn’t always
possible because most of the students would leave the system before they met the
standard of proficiency. Students who do not have literacy in their own language take
even longer to develop proficiency in English, Della explains:

They don’t have enough time to get all they need before they are forced into a
position where they are forced out of high school.... You know it takes five to
seven years or probably longer than that for these students because many of them
are not literate in their first language or they are not very literate in their first
language. I mean, we don’t have a cut-off English test that we would say, “No,
you can’t graduate because you’re not this proficient in English.”

The teachers who work with the child’s first language in a bilingual educational
approach, stress first language assessment and proficiency as important factors in student
achievement. Kim explains the importance of Native language literacy. She explains that
her district can only offer literacy support in certain languages. “We don’t have a teacher
that speaks Hmong” she explains and many schools in her district have up to 34
languages. She believes that the lack of home language support has a negative impact of
the child’s English language proficiency. “It’s definitely going to impact the success in
school,” she states.

Sharla who teaches a Native American language in a reservation school expresses
English language proficiency in terms of her philosophy of education which is formed by
her cultural values. Sharla is the most verbal about emphasizing the importance of
proficiency in the home language along with English. Proficiency is not just a set of
skills that can be assessed on a pencil and paper test, according to Sharla. Her definition
of proficiency supports a social constructivist approach in that it included the importance
of using different language skills in different contexts. Sharla explains:
I would think that if they’re able to talk in both languages, dialogue in both languages in different settings, that they’re able to do it in the right time and place…. and they’re able to be understood and get what they went there for.

Sharla explains that to her, the test of whether students are proficient or not in a language is whether “they’re able to be understood and get what they went there for!” She explains that her ancestors were knowledgeable in a number of languages and used those languages to navigate successfully in different cultural worlds. “We had leaders that were able to talk with people in Congress, you know and meet with the different officials, and, I think that those were the ones that were able to talk the language fluently too.”

Overall, the participants’ responses perspectives on the point in which students could be categorized as proficient in English and no longer eligible for services or accommodations differed. These differences were related to their work situations, educational philosophies and experiences in teaching and assessing students. Whereas most agree that the academic English language proficiency needed in school is an important issue, many other factors are also part of the picture.

Summary of Theme One

To summarize Theme One, including the sub-themes, it is important to stress that the participants in the study use a variety of assessments to learn about and document their students’ development of English language proficiency. They believe in the need to use a variety of assessments. Many reported that all the pieces of information about the student can come together to form a picture. They stress that informal, authentic and performance assessment can reveal a more complete picture about a student. This type of assessment is more supportive of students from different cultures and can reveal students’
strengths, along with weaknesses. Self-assessment and home language assessment are also types of authentic assessment that supports the values, learning styles and background information that the student brings to the classroom.

The teachers have different understandings of attainment of English language proficiency. This causes them to communicate information about student progress and attainment in different ways. They are also influenced by the requirements of their schools and states, the student population they work with and the demands of their jobs. Another difference among the teachers was their articulation of proficiency and exiting criteria.

**Theme 2: Teachers Know What Their Students Can and Cannot Do**

Assessment involves knowledge about a student. It involves gathering information about students from various sources. Though judgment comes into play more with the evaluation of assessment results, thoughtful decision-making also is involved with choosing tools to assess and including information as part of assessment. Theme Two involves teachers’ knowledge and also includes a contradiction in that teachers may express the fact that they know their students, but their ability to articulate this knowledge varies. The theme is organized into two sub-themes. These two sub-themes reflect the discrepancy between knowing students, but articulating knowledge differently.

**Theme 2a. Teachers Clearly Express That They Know Their Students**

One piece of information that is consistent with all participants is that they feel that they know their students. They claim they know what the students can do and what they can’t do and they know when they have made progress in learning. Most of them
stressed the value of the knowledge that comes from personally knowing students and having a relationship with them.

"I think a teacher really needs to know their students very well in order to be able to assess," Nona states. She stresses the need to use portfolios and informal assessments with students along with standardized test scores. She also stresses communication. "If you are going to be looking at a model like that you have to have lots of communication between lots of different people about that kid."

Along with talking to others, teachers learn about their students by talking to them and their family members. Iva explains the importance of involving the student in the assessment process. Students can be very informative. "But one of the first things I do when a child comes from a different country, I want that child to share about their own culture, their own alphabet...." Iva explains. Nona tells me the story of her discussion with the father of one of her students and how the father was concerned about the child was relying too much on her tutor. "I would definitely take into account the comments he made," she said.

Many of the participants report that gaining an understanding of the students' home language and culture, family and other things along with just their work in school gives them a better understanding of the students. The things they learn through interaction with the families and understanding the students' cultures help in the assessment process. Hettie explains the value of communicating with the students' parents and home community:

I will tell you a big thing it does for me is, and I only know this because mainstream teachers that don't have this information look at the kids way different than I do. I look at them and I know what they're good at, I know if
they're smarter.... So sometimes I think by knowing, by talking to the parents and seeing how the parents see their kids and then, also, just knowing how the culture is and how the kids are quite often very responsible ... it kind of gives me a heads up on what the kids are already good at, it gives me a heads up on why they don’t do some things.

The participants report that that knowing what students can and cannot do enters into assessment. They used their knowledge of the student as a sort of scaffolding in instruction and state that it is also important to take that same support to the assessment environment. Nona, who has a strong belief in the involvement of students in their own instruction, explains this type of scaffolding in assessment. “I think we need to help them learn about the process of here’s what’s expected and, if you say this, then you you’ve only given part of the answer so you only have a rubric like say one on a score of zero to five.”

Theme2b. Teachers Articulate Their Knowledge of Students with Varied Clarity and Often through Stories

Assessment involves pulling together information on a student. Assessment leads to the interpretation of that information. Assessment is knowledge gained about a student. It involves looking at the information and using it to explain what a student can do and cannot do. It requires the ability to articulate what you know and don’t know about a student. As stated, all the participants reported that they knew their students. Theme 2a focuses on the finding that the participants feel they know their students’ command of language proficiency. They know what their students can do and cannot do. Theme 2b is the individual variance in how this knowledge was articulated. This concept of articulation is important and will become more significant in the discussion of the findings.
Teachers articulate how they know their students in different ways. Iva expresses how she weaves authentic assessment into her work with the students. “So, I do a lot of observing checklists, running records, guided reading, self-assessments.” She also uses self-assessment with students and explains that even though a student may not know much English, it is still important to use because “it doesn’t matter how much English that child knows, they all really know how much they know.” Iva explains that it is important for students to be able to self reflect and say to themselves “O. K. these are the things that I’m really good at and these are the things that I have to work on.”

Catherine responds to my question on assessment more informally by telling stories about students:

I have a student right now – this Brazilian boy who is so driven. He does not need to be in my room anymore. He loves coming here. I think a lot of it has to do with the other kids are at a weaker level and he feels powerful, he feels, he likes that. He likes to excel. So he feels he can teach them and help them and he still keeps up in the classroom so it doesn’t matter if he comes to me.

She explains that after years of working with kids she knows where they are at and when they are ready to move on through “gut feelings”. Other participants referenced gut feelings and knowledge from experience.

Teachers value the fact that teaching is a relationship and good assessment is based on forming a positive relationship. “I’m all about relationships – building relationships with kids and getting to know them personally. The assessments I used were informal because I knew my kids so well,” Viola states. She explains that trust is an important part of that relationship. “when they trust you and trust that they’re safe in that classroom and then they feel free to ask questions and show you that they know things and try their best…. Otherwise they shut down and they won’t even try anything. Then
you can’t assess.” Viola states the importance of trust in the assessment process stronger in the second interview.

I think that most kids will do just about anything for you if they trust you. They will work hard ... So, I say, “You know, I need you to work really hard on this test. It’s going to show how good our school is and how good you are”. Then they’re going to do that for me. They’re going to work at that.

Stories were often used by the participants to tell about how they knew students were making progress in English. Hettie tells the story of a student who was not that proficient in English, but was able to use the English he had to explain that he needed to use her computer and printer and access tax information from a website.

I’ll never forget the little fourth grade guy that came to my high school room and said, ‘Can I print a form off your computer?’ and I said, ‘Well, what do you want to do?’ ‘Wee, Mom wants me to type and print off a tax form,’ and he knew where to go and what form she needed and everything else, but they didn’t have a printer at home ... that’s plenty good for fourth grade!

Catherine also uses stories to explain how she monitors progress and proficiency. She explains how different students progress at different rates and her experience and knowledge, along with observation helps her to determine when to start exiting a student from services. Her story of her boy from Brazil illustrates this point. She had explained that even though he was a new student, he moved along quickly and didn’t need ESL support very long.

Several teachers refer to knowing about their students and their students’ needs through “gut feelings”. Catherine explains that she would “go on just expertise or just gut feelings that I have that comes from years of doing it.” When asked to explain what goes into those gut feelings, she struggles to describe the various pieces of information that go into her views. She gives examples from her work:
The type of errors you see in the regular American kid who’s only been exposed to English are different from the ones coming from the foreign (students), the ones who struggle....The kinds of mistakes an American would make are not going to be idiom use or word usage.

In contrast, Iva, who had recently completed her Master’s Degree in Education, clearly expressed assessment issues, using professional terminology. Despite the fact that Iva’s speech reflects that English is her second language, her conversations express an understanding of language and literacy assessment. Iva states:

So, I’m just telling you the true assessment is a spectrum of assessments, different kinds of assessments that tell you this is the child, this child is this kind of learner and why don’t we just take that approach. You know like multiple intelligences.

Though the participants value authentic assessment and what it can reveal about students, the data demonstrate that they are not always taking the next step to look at information closely. I asked what the information Learning Record revealed on students that is not revealed on the standardized tests. Sharla responded with a laugh and explained that they are not looking at the data very closely.

Now that part I would not know. I mean we haven’t really looked at what they are doing. I don’t think anybody has done any in depth study on what they’re doing in school and…. We haven’t gotten together and said, you know, let’s look at these, let’s look at what we’ve been doing to our kids, what we’re subjecting them to.

Della explains that the teachers are required to use the SOLOM test in her state, which is State III. The SOLOM, discussed in Chapter II, is an informal observational oral language matrix and is their state requirement for oral language proficiency. Della explains the teachers don’t necessarily use the information it provides.

I mean I think they go through and give their rankings and that’s the last they think about it.... I don’t think anyone ever says, ‘Well, what’s their SOLOM ranking if they’re considering moving them from one level to another.
The participants expressed dismay regarding emphasis on the state mandated tests in the assessments that are used with students. They found these tests limited in the information they provided. But, the tests are not without value, they reported. The state tests, along with norm-reference commercial tests, can also reveal information about a student and be used to develop programs, instructional plans and advocate for student needs.

*Summary of Theme 2*

In summary, all participants claimed that they know their students. All participants reported that they use formal test results in learning about and describing their students. They also reported using informal assessments. Some of participants were able to discuss assessment and what their students could do using professional terminology from the field of education and assessment. Others were not as articulate in their use of professional terminology. Some participants expressed their knowledge of student development and achievement through stories or descriptions of student performance. Their knowledge was expressed in language that was more descriptive and intuitive and justified by their years of experience.

*Theme 3: Teachers Bring a Variety of Backgrounds and Contexts to the Assessment Process When Looking at English Language Learners*

Many factors enter into the complex activity of assessing and evaluating students. Teachers who are working closely with students have a powerful role in assessment because of the knowledge they have and the role they play. Yet, teachers participate in assessment from the context of their job expectations, their education, their personal backgrounds and philosophies and the current demands of their positions. The
information, viewpoint and goals they bring to the assessment setting is influenced by a number of things. The participants in my research study look at students, interpret assessment information, and value assessments from their different contexts.

The participants in my study come from a variety of backgrounds. Table 3 provides an overview of different characteristics of the participants. Participants include teachers in urban and rural areas and reservation schools. They include teachers who work with other bilingual and ESL teachers. They include teachers who work in isolation. They include teachers who work at a specific grade level and others who work with all grade levels. Some have graduate degrees and others don’t have a degree beyond a Bachelor’s. Some are in the midst of their career and others are close to retirement. Though all participants have teaching backgrounds, several are currently serving as program coordinators. Two teachers are responsible for first language development in a bilingual situation. Several of the teachers have had experiences living and working in other countries. A couple of the teachers shared the same language and ethnic background of the students.

The participants in the study were chosen because of their diversity. Within the group there is range in the ethnic background of the students, size of districts and geographical area. The participants, themselves, have varied backgrounds. There is divergence in educational background, experiences and ethnicity.

There are many commonalities among the participants. They all share the fact that everyone has a wealth of experience in education. All participants had at least ten years of teaching experience and have had experience assessing students. Catherine
explains how her years of teaching and knowledge of ELLs gives her a better understanding of what her students need than a standardized language proficiency test.

I certainly would not depend only on that test if I were to recommend a kid for ESL services or not recommend him. I would also go on just expertise or gut feelings that I have that comes from years of doing it.

Catherine explains that a new person coming in would not necessarily have the information that she had from her experience. "But, for someone new coming, I think it's going to be difficult." The new person does not have the knowledge that she does which comes from experience, and may not find it as easy as Catherine does to make assessment decisions.

All participants in the study share that they met their respective state requirements for licensure in education. Because the states varied in specific requirements for English as a second language teaching or bilingual education, not all participants had coursework in this area.

The participants also share a love of their work. A common code, which developed into a pattern, is this enjoyment of their profession. Consistently I would hear from the participants about the passion and love they have for their students and their work. "It's fun. I really, really like it," Kim states. Della, who is close to retiring, states, "It's just been a really rewarding career for me." Their enthusiasm for their work is also displayed in their commitment to their students. "Yeah, it's been fun. I tell you what? Every single student I have keeps me on my toes and I'm grateful for that. I never have a boring day. It's never boring," Hettie explains.

Another pattern from the data that relates to Theme 3 is the experience of being in a multicultural setting. Several of the participants had taught or traveled overseas.
Several are from minority backgrounds, themselves. These multicultural experiences influenced the participants and could be a factor in how they viewed ELLs, assessment and language proficiency. Those that did not have the multicultural experience had made efforts to participate in cultural activities and learn about their students' heritage.

The teachers who lived in other cultures feel that their experiences gave them greater understanding of the students they work with and helped them with assessment. For example, Kim states:

I've been to the countries they lived in and I've been a minority in a place and my first language was not the language of the country, so I've had a lot of the same kind of feelings in some respect.

Iva emphasizes the fact that she has been a refugee, just as many of her students are. Iva explains that she went to the university in her country and got a job in elementary education, but was not able to continue her career because of war.

But then in 1992 the war happened in my country and it was just, it was just a struggle to survive and I had a really hard time and then the first chance that I could take, my husband and we actually fled from Bosnia.

Iva explains how she brings her understanding of being a refugee and new immigrant to her work situation in teaching and assessing students.

Amy, who is not from the Native American background of her students, stresses that her experience on reservations gave her an understanding of the students' needs. "The last fifteen years I've been on the reservations," she states and names several of the Indian tribes in State I where she has worked. Sharla's Native American background impacts the viewpoint she brought to ELLs and assessment. She is related to many of her students and knew the families of the others. She understands many of the cultural issues the students bring to the testing environment, impacting their ability to succeed on a test.
that may not be culturally appropriate. She also values proficiency in the child’s first
language and culture, along with English language proficiency.

Another common area in their backgrounds that the participants brought to the
context of assessment is finding that all are committed to their jobs. Most of the
participants express words that confirmed the fact that they see their job as more than just
instruction. “All ESL teachers across the country realize that there’s so much more
involved than just teaching,” Catherine says. “You’re helping the families. You’re
helping them become acculturated.”

Hettie also explains the varied role of a rural ESL teacher. She explains with a
laugh:

I do a lot, especially after school, going out to parents’ homes and talking about
different things, like absenteeism in the high school and that kind of thing... a
little social work, a little counseling, a little... yeah, we keep busy!

A majority of the participants interviewed appear to be spread thin in their work
load. This concept arose with the different participants in different forms. The majority
of them expressed a strong dedication and commitment to their students. They stated that
they knew things about their students that are not revealed on tests but time was a factor
in their ability to develop and look at assessment information more carefully. They
shared with me schedules that were packed with students, responsibilities and duties.

Also, many of the participants lack the sort of security in their jobs that most
mainstream classroom teachers experience. They may be supported by “soft” or unsure
funding. They may be the only one in the district. They may be the first one in their
position. They may have a teaching area that is small and out of the way and a limited
budget for materials. Hettie explains that it took a while for her school district to
recognize the need and budget for ELL instruction. She explained to me that she had worked with the Laotian refugee population as a volunteer tutor for many years. "We had kids for a long time, but the school finally felt financially able to have an ESL teacher so I applied for the job and got it and so I have been doing it for thirteen years I think." Hettie explained to me outside the formal interview that the ESL program is still considered a new program and not every believes that is important. In a small town and small school district, school budget issues become everyone’s concern.

The teachers interviewed in the urban area did not have the same isolation and funding concerns as the more rural teachers, but they were in inner city schools that had related issues of instability, limited funding and marginalization. Marginalization could be seen as the relegation to an unimportant or powerless position within a society of group. Teachers on reservation schools also have similar issues of marginalization. All the participants in the study worked with a population of students that are considered minorities in a number of ways, including language proficiency, race, and socio-economic status. This fact could tend to reflect on them since the students were not always accepted by the mainstream system.

Codes such as "heavy caseload", "serves many grade levels", "wears many hats" "high student need", "one ESL teacher", and other related codes support that fact that the teachers have demanding jobs. Iva teaches in a large district in a rural state. She described the challenges of working in a program with a limited budget when she talked about her job:

We have a little over 300 LEP labeled kids which doesn’t give us a lot in funding, so we have just two ELL teachers and one ELL para who is shared between three schools.... And it’s so hard because they are all needy and then need attention...
Such a demanding job would not support the sort of time necessary to assess students in a reflective, holistic manner, and document that assessment information for others. Hettie doesn’t have the number of low level students, but the fact that she serves all the students in school district challenges her ability to provide the attention that some of the students need. Hettie explains that there used to be two ESL teachers to serve the students:

We used to have two. Then that person had, due to staff crunches, had to be moved into the library and teach some English classes and so now we have me and then we have a Lao paraprofessional.

Hettie explains that it is important to her to see all her students daily. It is a challenge, both physically and cognitively to work with students from all grade levels in the same day. She not only travels, but works with students at widely different developmental stages in one day. The value of meeting with students on a daily basis and being available for the classroom teachers is important to her.

I have thought, oh, maybe I should have a high school day and this and that, and I find that, in my experience it’s just great to touch bases everyday; to not get too far behind, for them to know that you’re there everyday. It might not be a long period of time, but you’re there everyday, for teachers to be able to communicate better with you too, because they know that they can catch you.

Though Hettie believes in the importance of seeing her students everyday, she doesn’t feel that she can meet all their needs very well because she has so little time. She laughs as she explains the challenges of being the only ESL teacher:

And one of my difficulties is having so little time to spend on them because of being spread out so thin. If you could have them, you know for English class, for an hour absolutely every single day.... But what I get is little snippets of time usually, you know, kind of like “propping up” time.
Hettie is spread so thin that she hardly has time to prepare for the next activity. Focused attention on authentic assessment and time to document that assessment and have a voice in decision-making is difficult, she states.

The codes "many hats" and "multiple tasks" came up with a number of participants. Several had very challenging jobs with a variety of tasks. Assessment of English Language Learners may have been only a small part of their job. Nona explains her job to me, "Yes, but I'm also responsible for all the testing and I'm maintaining the test scores and tracking all of the limited English proficient students. It's a huge job". Heavy caseloads are related to other issues ESL and bilingual teachers' experience, such as the insecurity of funding and the fact that it is not universally accepted, particularly in rural states, that ELLs have a right to specialized services and appropriately trained teachers.

Many ESL and bilingual teachers serve as paraprofessionals before they are recognized as teachers and given status as a teacher. The role, authority and recognition given to a teacher is quite different from a paraprofessional. Catherine, an ESL teacher in a mid-sized district in a State I, discusses the challenges she experienced in being recognized as a qualified professional with the same status as other teachers.

I started out as an aide in the school, just helping the foreign children get by. I was really teaching ESL to children, but I was being paid as an aide for about six years - no five years. Then, I got certified because I had a Master's but I wasn't certified to teach in the school system in (State I). I could just teach in the university level or be an aide and be underpaid. Then, I got certified and then I taught for a year with a certified contract, which was a huge difference in salary.

These issues which relate to a lack of voice, uncertain funding and status in the school support a sense of marginalization. Marginalization can be considered to be an
uncertain position in one lacks the status and recognition of groups that are part of the mainstream. Lack of voice and marginalization of the teachers is related to the students they teach. ELLs who come from a minority background and historically have been neglected or in mainstream educational programs can also be considered a marginalized group.

Summary of Theme 3

In summary, the teachers in the study brought different backgrounds to the process of assessment. The context in which they viewed assessment could have an impact on their role in assessment, an issue that is developed in Theme 4 and interpreted to a greater extent in the discussion of results. Overall, the participants expressed great enjoyment of their work and had wonderful tools to bring to their jobs, whether it was education or experiences in other countries. They all were impacted by heavy demands on their time and ability to complete the requirements of the job.

Theme 4: The Relationship Between Teacher Developed Assessment Information and Large Scale Assessment and Accountability is Multifaceted and Mixed

All the data from the different collection procedures, including the questionnaire, interviews and focus groups sought information about the participants' use of various assessments, participation in the large scale tests and viewpoints on the evaluative nature of the different assessments. Both positive and negative attitudes were expressed. The participants in the study are mixed in their understanding and involvement in the state assessment and accountability systems and requirements. They also have mixed feelings about the state tests, federal assessment and accountability requirements. They do not
feel they have a voice in the decisions made and the accountability systems that
document the work that they are doing with their students.

These data support the finding that the relationship participants have between
informal and teacher developed assessment information and that gained from the large
scale assessments is multifaceted and varied. The idea that these educators do not have
an understanding of their own state requirements for assessment was reflected in the
information reported in Table 19. Table 19 documents this lack of understanding. Three
of the participants admit to being not very knowledgeable about their state English
language proficiency assessment, according to the questionnaire.

Table 19. Knowledge of Legislation. The No Child Left Behind Legislation requires that
states and school districts develop new English language proficiency assessments based
on state standards and state definitions of proficiency. How knowledgeable about your
state’s English language proficiency assessment program are you? (n = 11).

<table>
<thead>
<tr>
<th>Not at all Knowledgeable</th>
<th>Not Very Knowledgeable</th>
<th>Neutral</th>
<th>Somewhat Knowledgeable</th>
<th>Very Knowledgeable</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Please Explain
• TEAE
• State II ELL Advisory Committee
• I know we do the TEAE test, but am not sure if it is a state, national or district
test. We also do the SOLOM.

The interviews support the data and documents that some participants have a
greater understanding of their state assessment requirements. The responses vary when
discussing their state requirements for assessment. Though the focus of the questions in
the interviews was English language proficiency assessment, the discussion typically
moved to the academic content tests. Participants were far more familiar with the
accountability terms under academic achievement, such as AYP or school improvement status. They were not as familiar with the counterpart of AYP for English language proficiency assessment, which are the AMAOs. Kim expresses her thoughts on NCLB requirements as determined in her state:

I am familiar with the fact that we’re measured by the state test – the English Language Learners are measured with the same criteria as the others and are often the reason a school will end up on AYP. But, other than that, I don’t, I’m not totally knowledgeable about how No Child Left Behind impacts ELL children.

The responses on the teachers’ participation and involvement with state accountability systems are also mixed. This participation would involve the development of items for the test, setting standards, participating in training or serving in an advisory capacity. Two have not been involved at all with two being neutral on the subject. Table 11, available on page 148 documents this mixed involvement. The written responses for this questionnaire item express the difference of experiences.

Some individuals had a great deal of experience with their state English language proficiency test development, such as serving on development committees, advisory committees and assisting with the training. Others felt that their experience as an administrator assisted them in understanding the state assessments.

The interviews provided data that expanded upon these questionnaire responses related to the participants’ involvement with their state tests and accountability systems. Catherine, in State I explains her assistance with her state’s development of an English language proficiency assessment as part of a consortium of states. She explains how teachers from the member consortium states worked with the test development vendor to write items for the test. “We all, from 11 different states, including (State II) I think, we
all went to Salt Lake and helped write the test. Measured Progress set it up.... Then they sent it here to be piloted”.

Amy, who is also in State I, expresses frustration related to her experience in implementing the state English language proficiency test. Amy states:

Well, all of us are dealing with [frustration], and the state’s frustrated too, because they have an almost impossible job to try to pull this off and, you know, starting at the beginning with no test and no experience, and so we’re all just waiting for more answers that they don’t have.

When asked how helpful the newly developed large scale English language proficiency tests were, participants had mixed answers on the questionnaire. Nearly half of the teachers found it not very useful or were neutral on the subject. Table 21 documents this difference of opinion.

Table 20. Progress Report Information. – The No Child Left Behind Legislation requires that states report on progress students made toward English language proficiency objectives based on the state English language proficiency assessment. Is this information helpful to you? (n = 11).

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Please Explain

• Yes, now our kids count and can’t be ignored.

Catherine explains more about the issues related to the new tests when she continues on with her discussion of the test she helped develop in State I. She explains that she didn’t like the product once the vendor put the test together and sent it to them for piloting. Catherine states:
I already smelled that it was wrong. Then they produced the real test but then we had to wait two years to do it because there was some glitch. I don’t know what happened. Then we just gave it last fall. I’m not happy with it. It is not a good test. It’s too easy, basically.

The participants were more consistent in the questionnaire regarding what they would like to see from their state English language proficiency assessments. They would like to see overall information on students’ proficiency in English, along with breakdowns of how the students performed in the sub-domains of reading, writing, speaking and listening. They would also like to see how students developed from the previous year. There was hope expressed in the interviews that the English language proficiency test would be more beneficial to them. Catherine explains that her state is going to work on improving their test:

Yeah, we are going to work on it. I mean no test stays the way it is. You do it and you make changes to it and so on and so I would assume that after a couple years, or whatever, it’s going to become very usable, user friendly.

Wanda is hopeful for the new English language proficiency test in her state, explaining that:

I think that if this is a true measure of their proficiency then we’re going to have some eye opening results come summer and we’re going to have students who we didn’t realize were as low as what they are and we’re going to need to have more ELL staff to provide the services for them.

Iva is also hopeful for the new test her state adopting. “It’s really appreciated because now we are adopting, implementing, the WIDA ACCESS assessment for English language proficiency which is really good compared to what we used before.”

The participants express more dissatisfaction over the academic achievement requirements and accountability than the English language proficiency accountability requirements. This dissatisfaction was expressed mostly in the interviews and focus
groups than the questionnaire. My focus in the interviews was in English language proficiency assessment. In the majority of cases the discussion usually ended up with the academic achievement assessment requirements.

"Frustration" became a code since it is a word consistently used by many of the participants when referring to the achievement tests. Della explains that she is impacted "quite a bit actually" by the federal requirements related to the No Child Left Behind Act in her role as a coordinator. "And in the paperwork that we have to do and the frustrations of that whole thing trying to meet the annual yearly progress goals with a population that will never meet the annual yearly progress goals."

Kim sees the expectation for English Language Learners to make adequate yearly progress the same as the other students as not fair. Others were also dissatisfied with the fact that this group had to meet the same academic goals as other students.

Sharla discusses the fact that the Native American heritage is not represented in the content standards that students are assessed against in the state academic achievement assessment program and the fact that most of the school districts that do not make adequate yearly progress (AYP) are Native American. "Most of the schools that didn't make AYP were the Native American. So there are a lot of concerns that we've become stereotyped."

Nona questions the importance of meeting the goals required for AYP in referencing a speaker she had heard at a conference that morning. "I think part of assessment that we miss on the standardized test is that critical thinking part, you know, being able to analyze something meaningfully." I interviewed Nona when we were both attending a conference in her state. Our interview conversation was influenced by the
key note speaker, James Banks. After his opening speech, Banks had responded to a question that asked his opinion of the current federal legislation. He had responded that unconscionable things have been done by highly literate societies.

Nona builds on Banks’ reference to the fact that the Holocaust took place in a literate, well-educated society, when explaining her thoughts on the accountability system that reports school districts in the state according to the progress their children make in math and reading scores on the state test. She states:

You know, whereas you might be able to take a very literate, well-educated, very intelligent person and, they’re the Hitlers of the future,” echoing Bank’ emphasis on a curriculum and standards that emphasize values and human relationships along with strong literacy and math skills.

The participants discussed the pros and cons of academic achievement assessment requirements. They explain that they understand that ELLs are now included in the report card or picture of how a school district is doing. The participants don’t always agree that it was a fair picture of how ELLs were achieving against state academic standards, but they did appreciate the requirement that ELLs participated along with other students. Iva explains that testing is part of our world and it is important for the students she teaches to be part of that world.

“You know what?” Iva asks.

This is the thing. This is how I take it as a teacher. We live in the world of the tests. Where ever you go. If you apply for a scholarship. If you want to apply to go to a good school. Where ever you go you will have some kind of test.

Viola is also pragmatic about it. “Even if I’m not agreeing with the measurement, at least my kids are being measured! It’s part of being on the map! These kids exist!” Viola laughs as she explains this. “I’m not a belligerent activist kind of person ... I guess
in this case I am!” Hettie reflects a similar position when she says “I think the state assessments woke up many districts and I give them (the tests) credit for that.”

The participants had different experiences in the development of their state English language proficiency tests. Since their experiences varied, from non-involvement in any aspect, to involvement in the creation of items and setting of cut scores and standards, they have different viewpoints on the assessments.

The issue that the teachers have had differing experiences in the development of their state tests and standards supports the finding that their relationship in the accountability system is mixed. They express a variety of opinions and have mixed feelings about the tests and the use of their own assessment information in accountability also confirms the idea that their relationship with the system is ambiguous. A number of codes and notes show that there was more pressure and support in the school districts for teachers to give attention to standardized test results than assessment that comes from them.

An example of the diminished support for assessment that is developed by teachers is provided from the interview with Sharla, who discussed the Learning Record used by the BIA schools. Sharla explains that the Learning Record is a portfolio of student work that involves student selection and teacher discussion. The strengths of the Learning Record, according to Sharla, is that it allows all the work of the student to have value. All the students have an opportunity to demonstrate their work. “We still got work out of a student that wasn’t doing well. There was still a grading system,” she explains.

She explains that the collaboration and the sharing of the folders is also a strength. “It was good for teachers to see that (if) they had a student that they weren’t able to reach
they could see that someone else was able to reach him because all of the teachers would have a folder.” She further elaborates on the idea that the Learning Record emphasized what a student could do. “I thought that was a good thing, instead of saying that they don’t work for me so they must not work for anybody else – adding on these labels. With this Learning Record you could see they were workers.”

The state academic achievement test required by the federal legislation that generates the AYP results has taken a greater focus. Sharia explains that teachers are still required to use the Learning Record “as far as I know, we were mandated to do the Learning Record and nobody undid the mandate.” Now the support is no longer there. The scheduled time to get together to discuss the student work no longer exists. Sharia explains “It’s still ongoing, but it’s not as organized I guess, where everybody met together and exchanged papers and looked at students’ progress in different areas.” The professional development they participated in to support the implementation also is gone.

Sharia explains, “Yeah, we had somebody come in and do training and come back and do the moderations and take it (the moderations) out to California and that part is not going on.” Sharia believes that the teachers and students have no voice in the current system. She explains that she doesn’t believe that paper and pencil tests are the best way to assess student knowledge and teachers often know that students can do things even though they don’t perform well on the test. There is no room on the test papers for the teacher information.

I think with the testing (a child) just becomes a number and there’s no place for teachers to put on the test or to (tell) the testing company or to the state or wherever these test scores go that they (the students) do know how to do this stuff. You now we can’t do attachments.

204
In discussing authentic assessment, Tierney’s principles of assessment and state accountability systems in the online chats, teachers express support for authentic assessment and wished that the philosophy could be incorporated into the accountability systems. Iva explains she supports Tierney’s ideas and believes assessment should be “developmental, suggestive and individualized.” She does not think that these concepts are compatible with the large scale state tests and accountability. The other participants like Iva’s description of assessment and agree with her. “These principles are great to read just to ‘teach’ the teacher,” Hettie states. The participants do not see these features expressed in Tierney’s principles in the state and federal requirements. “Nobody had read Tierney in the government,” Hettie jokes.

When asked, the participants do not have concrete answers on what relationship they would like to see between teacher developed assessments and teacher voices and large scale assessments. They are not sure though, how the voice of the teacher, the student and the student’s cultural community can become part of the accountability picture. They do strongly express that the voice does need to be included.

The focus group data included the most specific comments in the area of teacher voice and state and federal accountability. “The policymakers need to listen to teachers, the people in the trenches,” Sharla states. Iva recommends that policy makers should design assessment that is not just based on test scores and gives a more holistic picture of the child. She explains that they should:

design an assessment from which our students can benefit and which will drive our instruction. Let’s create the assessment which will give us a true picture of our students as learners - like you all know they are so smart, but the numbers are not showing that true fact.
Hettie provides a summary comment in the focus group to the discussion on the problem of including teacher voices and information in large scale assessment and accountability. "Alleluia and Amen to all of you! Wish I could have cured this in one Blackboard chat!" Hettie states.

Iva ends the focus group discussion on a positive note which expresses an enthusiasm and hope that I found in many of the participants. "Who knows maybe our dream will come true one day."

The findings that support the fourth theme addresses voice, involvement and relationship. There is frustration and concern with current large scale tests and accountability. This frustration and concern includes both the English language proficiency assessments and the academic content assessments. There is hope for improvement. There is both anticipation and resignation expressed over the accountability systems – particularly the content area tests that have greater impact. Participants expressed a hope that even though there was a punitive nature to the reporting of these results, there would be more recognition of the needs of ELLs because of the accountability systems. There is a sense of resignation in that the requirements are unavoidable. Testing is part of the educational environment.

Some of the participants have been involved in the development of English language proficiency assessments and accountability. The participants had less involvement with the academic achievement assessment programs. The question of whether their voice or the voice of the students and families they represent is included in large scale assessment and accountability decisions is a complicated issue that cannot easily be answered. On the whole the participants did not feel that the federal
requirements and state accountability systems which included the academic achievement and English language proficiency assessment programs included teacher and student voices. This lack of voice did not keep them from a commitment to their work and attempting to find ways to advocate for their students.

Summary

Overall the data reveal the variety of assessments that are used to assess English Language Learners. They show that teachers bring their professional knowledge and experiences to issue of assessment. It can be seen from the data that though teachers report that they know what students can do and cannot do, but this knowledge is articulated differently. Ultimately, the data documents that when the relationship between teacher developed assessments and large scale assessment and accountability is examined, the issues become even more complex. These findings will be further explored in Chapter V as they are interpreted and assertions are developed.
CHAPTER V
ASSERTIONS AND CONCLUSIONS

Introduction

"Don't walk behind me; I may not lead. Don't walk in front of me: I may not follow. Walk beside me that we may be as one." (Zona, 1994)

This Native American proverb addresses the importance of working together. It served as guide for me while summarizing my research study. A number of ideas surfaced in my data. One of the ideas involves the interest of teachers in having a voice, accountability, and working together with others in efforts to teach children and evaluate their growth. This concept of working together is an important one and significant to my findings and assertions.

In this final chapter I will provide an interpretation of my findings and themes. I will make several assertions and conclusions. I will summarize my discussion with recommendations for action and further research. Woven throughout the assertions and recommendations is the value of walking side by side as important decisions are made concerning children and young people in our nation’s classrooms.

I started this study in an effort to address a research problem that involved the various policies and practices in assessing students from language minority backgrounds. The literature showed a disconnect between theory and practice and between large scale, high stakes assessment and classroom assessment. With a theoretical basis that supports assessment information closest to the student, I focused on teachers’ practices and
I developed three questions to serve as a guide in my exploration of the issues. I used a mixed method design, based on naturalistic research theory and qualitative methods. I developed research questions to serve as my guide in exploring the research problem. These questions include:

1. What assessments are English as a Second Language (ESL), English Language Learner (ELL), and bilingual teachers using to document the development of English language proficiency of English Language Learners?

2. What do the teachers say the assessments reveal about the students?

3. What is the relationship between teacher developed assessment information and large-scale assessment data and accountability?

After collecting data through four different collection activities and analyzing it, several themes emerged. The themes are:

Theme 1: Teachers use a variety of assessments to understand English Language Learners.

1a. Teachers use formal and informal assessments to document student learning and progress towards proficiency.

1b. Teachers claim that both formal and informal assessment can reveal important information about students.

1c. Teachers’ perspectives on English language proficiency and program exit criteria depend on contextual factors.

Theme 2: Teachers know what their students can and cannot do.

2a. Teachers clearly express that they know their students.
2b. Teachers articulate their knowledge of students with a clear sense of knowing, and often use stories to explain.

Theme 3: Teachers bring a variety of backgrounds and contexts to the assessment process when assessing English Language Learners.

Theme 4: The relationship between teacher developed assessment information and large scale assessment and accountability is multifaceted and mixed.

As discussed in Chapter IV, though there appears to be an ordinary quality to each one of the research findings and themes, the findings are significant. There are commonalities in the findings. There are also paradoxes and contradictions. The participants claim that they know their students and yet they do not articulate that information clearly. The participants express a belief in informal assessment, and yet don't always clearly express how they use informal assessment. This contradictory quality, which can be seen in the themes, is also found in my interpretations and assertions. The interpretations also reflect the conflicting nature of education today in which there are so many voices all claiming to have the answers to assessing student learning and achievement.

In reflecting on my findings, I return to the perspective of post-modernism in which “truth” is often seen to be made, not found, and reality can be “socially constructed” (Anderson, 1994, p. 8). One’s understanding of learning, assessment, accountability and roles of teachers can differ according to one’s viewpoint and perspective.
My data demonstrate the contradictions between the perspective of teachers and policymakers. Both teachers and policymakers have different ideas about how to find out whether students are performing and schools are doing their job.

My data also show contradictions and different understandings of student learning and achievement among the teachers. Each has a different understanding of determining English language proficiency. There are also contradictions between the expectations of educators and the community that represents the students. ELLs often come from minority groups that may have different values from the mainstream culture. Life goals that impact learning may be different. Also, the literature has demonstrated that learning and language theory vary. Educational goals and policies may be based on theories of learning that conflict.

These contradictions and multiple meanings are inherent in a naturalistic study that allows for a social constructivist and postmodern world view.

A postmodern approach forgoes the search for true fixed meanings and emphasizes descriptive nuances, differences, and paradoxes. There is a change from a substantial to a relational concept of meaning, with a move from the modern search for the one true and real meaning to a relational unfolding of meanings. Different interpreters constructing different meanings of an interview story is then not a weakness, but a strength of the interview method. Meanings and numbers are constructions of a social reality. (Kvale, 1996, p. 226)

My interpretations reflect this multiplicity of meaning. My data can be interpreted in a number of ways. My themes are purposely general, allowing for different interpretations. Yet, within the contradictions and multiple meanings, there are specific viewpoints that must be considered because of their strength and authority. My assertions will reflect my findings and themes, but will also encompass theory, literature
and my own understandings. I will also use the data and literature to document reliability and trustworthiness of the assertions I make.

My research study took me on a journey in more ways than one. I traveled to visit with my participants. I journeyed through three upper Midwest states in my research. I traveled east and west and north and south. Some of the journeys involved not only the visits to classrooms where my participants work, but I also saw some of their communities. I spent the day with one of the participants, Sharla, as part of my interview. We stopped to visit a cemetery where her ancestors, who included famous Native American scouts and chiefs, were buried. As she shared her history with me I understood who she was better and how her background and belief system helped form her teaching philosophy.

I traveled figuratively as a researcher. As I had discussed in Chapter III on my research procedures, I journeyed with my participants, “wandering together” (Kvale, p. 4) in professional conversations, and traveled together on a “journey that leads to a tale to be told upon returning home” (Kvale, p. 4).

I also traveled through the exploration of theory, research and literature on my subject. I began with thoughts and concerns. I began with a compelling issue that has created national attention. I took this issue and developed a conceptual framework and research design. I reviewed literature to support the framework and draft research questions. I also continued to review literature and theory as I analyzed my data. I developed themes that emerged from the codes, categories and patterns of the data. From these themes I traveled further to make assertions, based on my findings.
Assertions

The completion of my journey takes place as I reflected on the findings and make interpretations and assertions. In one sense, I have come full circle in that I approach my participants again and interpret their conversations and responses to my questions. I experience again the first sense of awe and excitement I had when I started my research.

The journey has also taken me back to the original conceptual framework of my study. This framework includes the theoretical foundation and related research found in reviews of literature. The research study is based on a social constructivist understanding of language and learning that emphasizes the interactive nature of learning and the important role of a child’s culture. The literature that addresses language acquisition from this perspective supports the complexity of language and the need to recognize not only academic language proficiency, but also the many domains and registers necessary to function successfully in the world today.

Along with theory and research, the framework of my research also involves the voices of those who contribute to the study. It includes my participants, those who served as consultants and my advisor. Myself as researcher must also be included as a voice in the research. Though I set my personal biases and points of view aside when I entered the field to gather data, thought and decision-making went into data analysis, assertions and interpretations.

This framework serves as a basis for interpretations and assertions. It is also serves to support a sense of trustworthiness, validity and reliability for the findings and assertions. Kvale explains the reconciliation of naturalistic research in postmodern perspective that allows for multiple meanings with the traditional research objectives of
validation, reliability and generalization. He describes an approach that does not necessarily reject the concepts of multiplicity in interpretations, but reconceptualizes them in forms that are relevant to qualitative methods (1996, p. 231).

Kvale returns to the research design, methods and quality of craftsmanship in determining validity. “The craftsmanship and credibility of the researcher becomes essential” (1996, p. 241). Ultimately, the findings themselves speak to validity, according to Kvale.

Ideally, the quality of the craftsmanship results in products with knowledge claims that are so powerful and convincing in their own right that they, so to say, carry the validation with them, like a strong piece of art. In such cases, the research procedures would be transparent and the results evident, and the conclusions of the study intrinsically convincing as true, beautiful, and good. Appeals to external certification, or official validity stamps of approval, then become secondary. Valid research would in this sense be research that makes questions of validity superfluous. (1996, p. 252)

I believe that the claims I make, based on my research findings, are powerful. Within the assortment of meanings that can be taken from the data, there are some observations that surface because they are compelling. They are compelling because they resonate as “true, beautiful and good” (Kvale, 1996, p. 252). The claims I make are based on the voices of the teachers who participated in my study and I offer them as one more perspective in the cacophony of voices in the debate on policies and practices in the assessment of ELLs.

Figure 8 illustrates the assertions that I developed based on my research findings. The assertions are placed in a landscape, similar to my research design in Figure 3 and my Research Themes and Findings in Figure 7. This landscape represents the research
Assertion 1:
Teachers of English Language Learners need to be supported in a task that takes courage and passion.

Assertion 2:
Teachers of English Language Learners need to be supported in navigating a course between the large scale tests for accountability and their own understanding of student achievement.

Assertion 3:
English language proficiency accountability systems and policy need to be developed with those who are being held accountable.

Assertion 4:
Teachers of English Language Learners need to be supported and scaffolded in articulating information on their students.

Assertion 5:
English language proficiency assessment needs to be based on a theoretical foundation that reflects that complexity of language.

Figure 8. Research Assertions.
journey, encompassing the understanding of interviews as conversations that unfold. The background landscape includes many of the issues that were apparent in the research design of Figure 1. They are overshadowed, though, by the assertions, in an effort to emphasize the power and importance of the claims that come from the data.

Assertion 1: Teachers of English Language Learners Need to be Supported in a Task That Takes Courage and Passion

I borrow from others in coming to the conclusion that teachers who work with ELLs need courage, strength and passion. Most specifically, I rely on Palmer who describes the incredible demands of the teaching profession in, *The Courage to Teach: Exploring the Inner Landscape of the Teacher's Life* (2007). Palmer discusses the integrity, self-reflection, sense of vocation and inner truth that are characteristic of a good teacher in much of his work (2007; 2003). Courage to Palmer is related to a teacher's ability to listen to the voice of the teacher within.

The teacher within is not the voice of conscience but of identity and integrity. It speaks not of what ought to be but of what is real for us, what is true. It says things like 'This is what fits you and this is what doesn't'; 'This is who you are and this is who you are not'; 'This is what gives you life and this is what kills your spirit – or makes you wish you were dead.' The teacher within stands guard at the gate of selfhood, warding off whatever insults our integrity and welcoming whatever affirms it. (2007, p. 32)

Palmer reaffirms this view of teachers and the ongoing need to support the teaching profession in maintaining a sense of commitment and integrity in the midst of current challenges in his foreword to the tenth anniversary edition to the book. He explains that this courage and integrity that he described when his book was first published are even more important today.

I was also right about the way inner work can help teachers connect with their students (thus aiding and abetting learning) and empower them to resist the forces
that threaten to undermine real teaching (of which NCLB is only the most recent example). (2007, p. xiii)

One finding that is consistent for all my participants is the passion and commitment they have for their work. Even though I did not ask if they liked their work, the issue surfaced in the first interview and was expressed in different forms in all of them. Kim, who was my first interviewee, brought up the subject, expressing her positive feelings with such enjoyment that took me a bit by surprise. “It’s fun. I really, really like it!” Kim explained in discussing her work, her students, her field of ELL, and bilingual education.

The appreciation for their work, along with their love for their students was expressed in some way by all the participants. They form positive and affirming relationships with their students and families. The stories they tell about their students express this connection. My field notes reflect the vibrancy in their voices and gestures when they talk about their students. The finding that teachers love their work and their students didn’t become a major theme since it did not directly relate to my research questions and problem. It does, indirectly, relate when I contrast it with the challenges all of the participants face in their jobs.

The paradox to the finding that the participants love their work and their students is that the participants also face significant challenges on the job. Nearly all participants addressed the issue of busy schedules, heavy caseloads and limited resources. These challenges could easily cause an individual to become demoralized and unhappy.

Iva’s summary of her day overwhelms me. “Mostly my day goes from 7:00 to 5:30 - 6:00,” she explains, listing all the schools and students that she interacts with.
everyday. Her day often ends with committee meetings in which she needs to advocate for ELLs in a community that doesn’t always recognize their needs. Iva explains that “We have to be bold as ELL teachers because you know this community needs to know whatever is needed to help those kids.” I admired Iva’s energy and her willingness to be “bold” for her students. She had a demanding schedule, limited resources and yet took the time and courage to advocate for her students.

I also admired Hettie’s willingness to speak up for her students with a school district that did not always understand their needs. Hettie faces some of the same issues related to misunderstandings of the needs of ELLs. She explains that she advocated for her students with the superintendent.

I’ve tried to tell the superintendent … I said, “Would you please just talk to me someday and I will tell you what I think we are wasting our time on and what really makes it so a kid can graduate from high school because we do so much busy work and so many worksheets.”

The difficulties that educators who advocate for ELLs is supported by studies that explore educator attitudes toward students who came to school with different language backgrounds. Walker, Shafer and Liams surveyed mainstream classroom teachers in a Midwest state school district with a growing ELL population (2004). They found that “the overall nature of teacher attitudes towards ELLs appears neutral to strongly negative in this study” (2004, p. 140). Some of the factors that contribute to the negative teacher attitudes include:

- Time and teacher burden; ELLs in the classroom are demanding on time,
- Lack of training; mainstream classroom teachers have little understanding of appropriate teaching strategies for ELLs,
• The influence of negative administrator attitudes,
• Malignant misnomers about effective ELL education, and ethnocentric bias (Walker, Shafer, & Iiams, 2004).

Whereas the focus of the research of Walker, et al. was primarily the mainstream educators, their findings do support the challenges ESL and bilingual teachers face in their jobs.

When we asked ELL teachers the interview question, “What do you see as the largest obstacles in implementing a quality ELL program in your school?” almost all of the responses included negative teacher attitudes. The ELL teachers repeatedly mentioned the unwillingness of many classroom teachers to make adaptations, or to have ELLs placed in their classroom. They mentioned the difficulty in finding classroom teachers who are interested in collaborative teaching, and the frustration of working with teachers who think ELL students should be the sole responsibility of the ELL teacher. (2004, p. 143)

Along with the job struggles and need to advocate, nearly all my research participants discussed the challenges related to working with a group of students that were a minority and did not make achievement goals. The participants expressed the frustration or unfairness of a system in which their students were targeted for not achieving at the same level as other students. Minugh, a reporter with the Sacramento Bee, describes the requirements of the law in looking at the dilemmas at the Will Rogers School in California face with their large ELL population.

The federal law holds schools accountable for their overall test scores as well as for the scores of groups typically underperforming students – like non-English speakers or the ethnic minorities. A whole school can be penalized if one group repeatedly falls short, like at Will Rogers. (2008, Sect. 4)

Kim, who works in one of the largest districts in her state, explains that because of the ELL population in her school, it becomes identified as a school that is not making adequate yearly progress under the Title I NCLB requirements. “And so we can become
an AYP school even if our English Language Learners are not all functioning at their grade level in English. That’s not fair.”

My participants’ courage, passion and commitment to their students becomes an issue related to the research because of the challenges of working in an educational environment with an accountability system that places a great deal of outside pressure on teachers. Teachers who work with minority students have the responsibility of helping students succeed on the high stakes large scale assessments that are used to determine whether schools are making progress.

Though opinion differs on the fairness of holding schools accountable for the progress of ELLs at the same rate as other students, it is clear that teachers feel a tremendous amount of pressure. This pressure can cause demoralization, discouragement and job resignations. Teachers who maintain commitment and passion in the midst of this pressure should be commended and supported.

My second assertion evolves from the first and addresses some of the issues related to the relationship of teachers and large scale assessment. It focuses on the relationship of teachers and accountability policies. It involves a pragmatic understanding of the need to exist in an educational environment where theories, practices and policies clash.

_Assertion 2: Teachers of English Language Learners Need to be Supported in Navigating a Course Between the Large Scale Tests for Accountability and Their Own Understanding of Student Achievement_

My research findings demonstrate my participants believe strongly in the assessments that they develop in the classroom. They have mixed attitudes towards the large scale assessments. They appear to have to navigate a course between the
accountability systems and their own belief in assessment and student learning. This helps them maintain their commitment to teaching and learning.

Teachers who participated in my study expressed mixed ideas about assessment and their relationship with large scale assessment and accountability systems. They believe in their own assessment information. They are often frustrated with the state tests because the information doesn’t reflect what they know about their students, along with other difficulties. They are also practical about the role of testing. In some cases, they see benefits. Along with an understanding of the benefits, they also discuss hope for improvement.

Overall, the participants express a belief in assessment that is holistic and student-centered. They valued informal and performance assessment. They believe in their own knowledge of what students can do and cannot do. They understand the influence of culture on students and how traditional testing situations conflict with cultural norms.

Amy explains how her students’ culture can negatively impact their success on a test:

So, they can’t take the tests that are written in academic English and really show what they can do. In their culture, too, they’re not as verbal in English. When I go to a teacher meeting the Crow speakers – they’re talking very fluently in Crow to each other, but then when they revert over to participating in English with the group, they don’t do it as much. They’re quiet, shy, reserved. I don’t know what it is – but they don’t participate as much.

The participants experience frustration because they see some of the conflict in an assessment system that doesn’t always reflect what their students can do. Amy also addresses some of the federal requirements for testing and standardization with mixed
feelings. "We’re just all frustrated, that’s for sure, in the state." She explains that the increased emphasis on testing has taken away more time from teaching.

Frustration with testing requirements is a pattern with the other participants, as I addressed in Chapter IV. I also discussed how the participants did not necessarily dismiss the use of tests altogether. Several gave examples of areas where standardized testing provide information on where a student’s language skills ranked against other students or against a standard. Therefore, they use the information from tests and accountability systems when it could benefit them. They are also realistic about testing. Iva believed strongly in informal assessment and describes the assessments she used with students, including self-assessment. She also explains the importance of teaching students how to take standardized tests:

We live in the world of tests; wherever you go, if you apply for a scholarship, if you want to apply to a good school, wherever you go you will have some kind of a test, even the teeny, tiny job.... We live in the world of tests because that’s how they measure our knowledge; multiple choice tests. My belief is that, it’s not my philosophy of teaching, but I have to teach my kids how to take tests.

Iva explains that within the world of tests it is important for a teacher to evaluate the information from the test. “Again, we have to make a distinction between the really good tests and the really bad tests.”

The participants also see the value for their students in the large scale assessments and accountability systems. Viola shares a concept that I found with several participants. “I am grateful to No Child Left Behind because my kids count now. They’re counting and people have to pay attention to them.” She explains that before the ELLs did not participate in the state academic achievement tests or if they participated and the score was not very good it was pulled out. Now, Viola explains, students can’t be excluded.
Viola states, “Even if I’m not agreeing with the measurement, at least my kids are being measured!” Emphasizing a positive approach Viola explains that the testing shows that “These kids exist!” Another expression that Viola used to show the positive effects of ELLs being included in the large scale accountability systems was the fact that “They were on the map.” Since ELLs historically were not included in state tests and there was not information on their academic performance, Viola and the other teachers see their inclusion as progress.

The participants express value in the large scale tests, but continue to use teacher developed assessments to learn about students. Viola discusses the assessment training that was taking place in her district and that more educators are using assessment. She states that “I’ll finish by saying assessment used to be a dirty word, I think, kind of scary, like something we have to do, but, I think that that’s changing and that’s good.”

Viola’s comment and some of the other participants’ views of assessment reflect a belief in spending more time with assessment because of its critical role in teaching and learning. Cobb explains:

As levels of accountability increase, teachers and administrators frequently lament the amount of time devoted to testing. They wonder why so much time is spent assessing student knowledge. I would like to propose what I consider to be a thoughtful suggestion to administrators and reading specialists for dealing with this dilemma: We need to spend more time assessing student knowledge.

(2005, p. 20)

Cobb explains that her recommendation to spend more time assessing students involves not only time on testing, but also time with informal authentic assessment. She stresses that educators learn the assessment language and become advocates for assessment (2005).
The courage and passion I saw in my participants help to support them in finding their way in the midst of testing requirements, numbers and public reporting of student success and failure. This navigation between their own assessments and the large scale assessments allows them to use the best of both and maintain their focus on student learning. It also allows them to maintain the passion and commitment they have for their work and their students.

*Assertion #3* *English Language Proficiency Accountability Systems and Policy Need to be Developed With Those Who are Being Held Accountable*

Della expresses her opinion on the assessment and accountability systems in education today, explaining that teachers know about students, “The policymakers need to listen to teachers – the people in the trenches – we taught them.” The other participants, as stated, also express dissatisfaction with accountability systems that were outwardly imposed. They express a trust in their own colleagues’ viewpoint on student achievement, rather than an external system or test. Viola explains the importance of using multiple assessments and teacher involvement in the high stakes decision of student proficiency. “No, a single measure can never be used to prove proficiency. I would trust a teacher’s judgment over a test score.”

The participants’ expression of frustration is stronger in reference to the academic achievement accountability systems rather than English language proficiency. As discussed, earlier, many of the conversations that began with the English language proficiency accountability system ended up to be discussions on academic achievement testing and accountability. The participants do not have much to say about the English language proficiency assessment goals – AMAOs. They do have a lot to say about AYP.
This difference could be because of the greater involvement of the teachers, themselves, in the English language proficiency assessments. They have more ownership in this program. The teachers in State I and State II, particularly, had more connection and involvement. It could also be because the consequences for not making achievement goals are greater in the academic achievement tests that all students took. English language proficiency assessment results do not make headlines to the same extent as the AYP issues. There is also a greater sense of disempowerment with the academic achievement accountability system.

The participants explain some of the importance of teacher involvement. Viola explains in a focus group why she trusts teachers' judgments over a test score.

We do depend on a single score on the test to tell us a lot — to get a feeling for where a student should be placed — on what level we can expect them to perform at in school. But if the teacher's findings are different — we may need to reconsider — to ask questions.

Others agree with her comments. They see value in the use of testing information, but they don't think that test scores should be the only thing to be used in accountability and high stakes decisions.

Having a voice in systems that impact us is a human need. Vella discusses the importance of engagement and involvement for adult learners. Though her principles are applied to adult education, they are relevant to systems in which adults work. Adults need to be respected, engaged and to feel as if they are part of the process (1994). In my literature review I addressed some studies which found that the current accountability systems have decreased teacher morale in some areas. McNeil and Coppola discuss the need to involve many in the development of educational policy (2006).
Policies are usually made far removed from those they govern, with the result that expert knowledge residing in the target organization or people is not incorporated into the making of the policy or into thinking through its possible implications. Research on policy impact needs to capture the voices of those affected not just because they are recipients of the policy but because they may have insights unavailable to the formal policy process; the power differential favors the policymakers, whereas actual knowledge differentials favor the professionals and families being affected. (2006, p. 683)

McNeil studied reform efforts in Texas and found accountability systems that are based on a system of controls, such as that imposed by NCLB, can have an opposite effect than expected (2005; 2006). This negative impact can involve lowering achievement and success of minorities. McNeil followed school districts that were impacted by the top down and standards based accountability reforms and found several factors contribute to this phenomenon. Along with the decrease in teacher morale, she cites the reduction in curriculum as teachers teach to a limited set of standards that will be assessed. Other issues, such as the increase of drop outs of minorities, cause the overall effect on minorities to be negative.

McNeil explains that the long-term effects of the standardized accountability system is very damaging. She states that “standardization widens educational inequalities and masks historical and persistent inequities. Standardization shifts both the control of schools and the official language of educational policy into a technical mode intended to divorce the public from the governance of public schools” (2000, p. 230).

Ingersoll (2007) finds that accountability systems that are based on control at the top do not always yield the results expected. He explains that an implicit viewpoint behind accountability systems that impose standardized measures, such as NCLB is that
greater control will result in higher student performance. Yet, what he calls the “tighten the ship perspective” can backfire (2007, pp. 21–22).

Ingersoll states that reform based on standardization and top down control suffers from three problems: The first problem is that the diagnosis is not correct. He explains that “the data show that the high degree of centralization in schools and lack of teacher control of their work – and not the opposite – often adversely affect how well schools function” (2007, p. 25). The second problem is that the reforms are unfair. He explains that that research demonstrates that teachers do not need the strict oversight that policymakers tell the public that they need.

Policymakers and reformers often question the caliber and quality of teachers, telling us time and again that teachers lack sufficient engagement, commitment, and accountability. However, the data suggest just the opposite – that teachers have an unusual degree of public service orientation and commitment and a relatively high “giving-to-getting” ratio, compared with those in other careers. The critics fail to appreciate the extent to which the teaching workforce is a source of human, social, and even financial capital in schools. (2007, p. 25)

Third, Ingersoll explains that accountability reforms often do not work. He explains that in denying teachers the very power and flexibility they need to do their jobs effectively, motivation is undermined and a valuable resource is squandered.

Having little say in terms, processes, and outcomes of their work, teachers may doubt they are doing worthwhile work – the very reason many of them came into the occupation in the first place – which may contribute to high rates of turnover. Consequently, accountability reforms may not only fail to solve the problems they seek to address, but actually end up making things worse. (2007, p. 25)

Barrentine and Stokes explain that despite the data that demonstrate that even though test scores may be increasing, student learning is not necessarily increasing, “educators are dealing with perceptions of changes in achievement in schools due to high stakes testing” (2005, p. 3). There is confusion and a lack of understanding of what real
learning involves and what the test scores and annual yearly progress reports mean. This statement made in 2005 is validated by recent reports that perceived increases in test scores may be misleading (Center on Education Policy, 2008; Institute for Educational Policy, 2008).

Along with the voice of the teachers in educational accountability, there needs to be the voice of the students and community. Educational policies that presume to set standards of achievement without the involvement of those who are impacted perpetuate policies reminiscent of colonialism and Indian boarding school policies. This point is particularly significant in a study that addresses a minority population, such as that of the ELL group. In many ways these students and their families are not only limited in English language proficiency, but also limited in having a choice in policies that influence their lives. This is not only because of their proficient language skills, but their rank in society as a minority group.

Frieire discusses the importance of including the voice of those who have been historically oppressed in educational systems (2005). Including this voice is important not only to those who have been oppressed, but also to those who have had the power and control. When both work together for liberation and freedom, society is transformed. Frieire explains that those who have been the minority, or as he considers it, oppressed, need to be part of the educational process that empowers them.

The pedagogy of the oppressed, animated by authentic, humanist (not humanitarian) generosity, presents itself as pedagogy of humankind. Pedagogy which begins with the egoistic interests of the oppressors (an egoism cloaked in the false generosity of paternalism) and makes of the oppressed the objects of its humanitarianism, itself maintains and embodies oppression. It is an instrument of the dehumanization. This is why, as we affirmed earlier, the pedagogy of the oppressed cannot be developed or practiced by the oppressors. It would be a
contradiction in terms if the oppressors not only defended but actually implemented a liberating education. (2005, p. 54)

Frieire emphasizes that educational policies developed and implemented by all involved have the potential of being liberating, authentic and humanistic. He differentiates between humanitarian policies which are reflective of paternalistic attitudes and humanistic policies that include the voice of the minority.

Nieto describes an educational reform system that encompasses Frieire’s theories of empowerment and Dewey’s emphasis on democracy in education. She stresses the importance of empowering students and critical thinking as part of reform efforts that lead to democratic societies.

School reform measures that have as their underlying focus both the empowerment of students and the creation of socially just learning environments are based on the view that critical reflection and analysis are fundamental to the development and maintenance of a democratic society. Without this perspective, learning can be defined as simply ‘banking education’ or as the depositing of knowledge in otherwise empty receptacles. But if we expect schools to be learning laboratories for democracy, where all students know that they are worthy and capable of learning and where they develop a social awareness and responsibility to their various communities, then classrooms and schools need to become just and empowering environments for all students. (1999, p. 174)

Assertion #4: Teachers of English Language Learners Need to be Supported and Scaffolded in Articulating Information on Their Students

One of the issues related to involving teachers in the assessment of the students they work with, is the need for teachers to have the professional skills to articulate assessment information. While I found that the teachers believe in their own assessments and claim that they knew student progress and achievement, their ability to articulate this information in a clear manner using professional terms varies. Catherine explains that
her knowledge of students has more merit than a test score in making high stakes decisions, such services.

I certainly would not depend only on that test if I were to recommend a kid for ESL services or not recommend him. I would also go on just expertise or just gut feelings that I have that comes from years of doing it (teaching ESL).

Others echoed Catherine’s belief in teacher judgment over test scores. Unfortunately, I found that Catherine struggled to express her understandings of students in vocabulary and terms that have common meanings in the field of educational assessment.

One’s first observation is that Catherine and the other participants’ lack of articulation meant that that they really did not know their students. If they cannot express this information with professional vocabulary and authority other than references to gut feelings, should they be considered credible? Along with the data that demonstrates a lack of articulation is the data that documents hesitancy when asked about specific assessment plans and criteria. Though teachers claim to believe in their own assessments and use informal assessment, the interview discussions often turned to the standardized test information.

Vygotsky stresses the relationship between thought and language, emphasizing the connection between one’s thought processes and the ability to put meaning into them through words and language. Both thought and language influence each other. “The connection between thought and word, however, is neither preformed nor constant” (Vygotsky, trans., 1986, p. 255).

Many might argue that, indeed, teachers either do not have the capacity to make assessment decisions on students or that it isn’t their role. The large scale tests that are
part of the high stakes accountability systems are based on the science of psychometrics, measurement and statistics. Popham compares teachers to another profession, medical doctors, in explaining that teachers should not have to create all the tools they need to function professionally. “Would you ask surgeons to construct their own scalpels before they operated?” (1997, p. 80).

I do believe Catherine knows her students. The data from the other teachers who expressed a trust in their colleagues’ information on students as more trustworthy than just test scores would also support my trust in her. The literature on social constructivist teaching and assessment would also support the teachers’ knowledge of her students.

I interpret this issue related to articulation to mean that teachers are not being supported in developing their own knowledge. I believe teachers can know their students and have the capacity to tell the public what their students’ capabilities and deficiencies are academically. Unfortunately, they are not being supported and affirmed in this ability. I rely again on Vygotsky and use his theory of the Zone of Proximal Development (ZPD) to discuss the need for teacher support. Vygotsky sees the ZPD of a child as the distance between the actual developmental levels of a child when working independently and the level of potential development when working under adult guidance or in collaboration with peers (1986).

Though this theory was developed with children in mind, I apply it to adults and use adult learning theory (Vella, 1994; Knowles, 1998) to state that adults are more successful when being mentored and when working with colleagues. Adults also need to be supported and scaffolded and will reach their potential more successfully with support. Teachers need to have opportunities to discuss and compare observations and conclusions.
about children with other teachers. “Teamwork is a principle of adult learning as well as an effective practice” (Vella, p. 149).

I take the model of Carini’s (2000, 2001) descriptive review of children to provide an example of educational assessment programs that are both collaborative and supportive. Carini guides teachers in her role as facilitator of the inquiry sessions in which children are discussed. She supports and scaffolds the learning, just as a teacher would support a learner in reaching his or her ZPD.

My role is to gather up the threads of the individual descriptions, delineating patterns where I find them, pointing out any questions that are forming, highlighting overlapping points of view and areas of difference. I try to keep these summaries brief during the meeting so that there is time for group members to respond to what they have heard in what their colleagues had to say. I write a fuller summary in the narrative notes. (Himey & Carini, 2000, pp. 197–198)

It is this process of facilitation that allows the teachers to express the knowledge they have about a student in a manner that is understood by others. The information comes from the teacher, but with the support from the facilitator, she or he is able to better articulate and create meaning to the knowledge within. I believe that the teachers of ELLs could also be more articulate if they were supported, affirmed, and guided in the appropriate expression of their knowledge.

It is also important to note that the guidance comes not only from the facilitator of the descriptive review inquiry groups, but also the colleagues.

The thing is, something significant happens when people commit to a sustained conversation around a shared topic of inquiry. That “something” exceeds the actual words spoken, and it can’t readily be replicated or summarized later for those who weren’t there. This is true of teachers’ groups and workshops and race dialogues circles and classrooms and so on. (Himley & Carini, 2000, p. 199)
Other educators have the capacity to scaffold and support each other. This emphasis on group interaction in the learning process is also consistent with progressivism and social constructive learning theory.

**Assertion #5: English Language Proficiency Assessment Needs to be Based on a Theoretical Foundation That Reflects the Complexity of Language and Learning**

The participants in my study learn about their students in a variety of ways. They learn about them through observation and interaction. They talk with other teachers. They talk to the students’ parents. They use their own assessments. They also use standardized assessment information. The participants tell about their students through stories and examples. Their manner of expressing information about students’ language development supports the complexity of language. It is also a very human activity that is not easily assessed through pencil and paper.

Several of the participants express the importance of a relationship and trust in assessment. They express the idea that students do better on classroom assessment because they are more comfortable in that environment. Kim states that “A classroom assessment can be an individual snapshot of the whole student; in a comfortable (hopefully) environment and they usually do better than on standardized tests.” Teachers who know students can better assess them. “When you spend time with a student, you’re more likely to be able to assess his/her proficiency level,” Wanda states.

Many of the participants’ comments reflect a social constructivist view of language development. Teachers who know their students, work with them everyday, learned their culture and their families, are able to see strengths in the students that were not necessarily reflected on the numerical test scores. Fayden (2005) discusses the role
of a social constructivist view of learning in children considered “at risk” for failing in school. She found that when she entered the lives of her students and used teaching methods that were supportive of social constructivism, the focus was on their strengths, not the deficits. She explains some of the reasons students are labeled “at risk” includes testing students fail because of its cultural incompatibility.

The point I want to make is this: The use of inadequate testing can cause inadvertent racism because it leads educators to view their charges as having shortages. Especially at the Kindergarten level, where standardized tests are perceived by the staff as determining how well the home environment has prepared the child for school, the reproach is readily placed on the culture and home life as having inadequately equipped the children. Therefore, not only is the child labeled at risk, but also he or she, presumable, is unfortunate enough to be born into a culture that has placed him there. (2005, p. 23)

Carini’s approach to assessment, which is based on observation and description, reflects this emphasis on children’s strengths. She explains the role of observation in assessing children:

The more I was in the close company of children, the more I observed, the more I was impressed that what was educationally consequential was how each of the children engaged the world, and the complexity and variety of that engagement. (2001, p. 184)

Carini explains the shortcomings of assessment systems that do not acknowledge this complexity:

What seemed lost from view is that classification, generalization, and the normative and statistical methodology more generally cannot accommodate or illuminate individualness and particularity. And cannot do so for the good and ample reason that these methodologies are not applicable to individuals or particularized experience. Categorization is a necessary condition for statistical manipulations of data to occur. Statistics and classification apply to groups and generalized phenomena and to distinctions among such groups and classes. I draw the inference that when individuality and the concrete particular is a dominant and important feature, as I suggest it is for anything human, the function of statistical methodologies is necessarily subordinate and limited. (2001, p. 185)
The need for relationships, trust and a caring environment for assessment to be effective was a common theme for my participants. As Viola states “all assessment is going to cause stress for kids.” She explains that “if they trust you, then they’re going to feel like, she’s in this with me. We’re in this together. I think you are going to get better results.”

Barrentine and Stokes explain that struggling students particularly can be “discouraged and defeated by high-pressure, high-stakes tests” (2005, p. 2) and assessment that is more holistic and based on the classroom can be more caring and supportive. “Classroom assessment involves a routine of caring relationships (Graves, 2002), well-chosen assessment practices aimed to understand the whole child, record-keeping, analysis, communication, and daily accountability” (Routman, 2003), (Barrentine & Stokes, 2005, p. 2).

Tierney’s Principles of Assessment available in Appendix B embody the learning and assessment theory the research participants, Fayden, Carini and other social constructivist researchers and practitioners discuss. They do not require excessive standardization, but allow for individual cultural differences. They also include the need for relationship, trust and caring. He explains:

These principles for assessment emanate from personal ideals and practice as much as theory and research – a mix of child-centered views of teaching, pluralistic and developmental views of children, constructivist views of knowing, and critical theoretical views of empowerment. (Tierney, 2005, p. 23)

These Principles which emphasize the humanness of the student and the knowledge of the teacher can be used to, not only provide information for classroom and public use, but also to empower the students and the community of the students.
Conclusions and Recommendations

Can policies be changed to include the voices of those who have not traditionally been included? Can teachers be supported in a manner that encourages strength, integrity and commitment, along with professional knowledge? Can children be assessed with assessment tools and systems that empower and affirm, rather than categorize and label? Can standards and rubrics developed to assess and determine the truth of who is making progress allow for the uniqueness and individuality of children?

These questions are both abstract and concrete. The answer can be “yes” to all of them. The practical reality of addressing the issues, though, is a different story. I would argue that just as Palmer addresses the courage of a teacher to maintain his or her convictions, we all must have courage and strive for what we believe in. Change starts at the individual level. Change also involves many people.

I make several recommendations for policy and research, based on my assertions. These recommendations will begin with those related to policy and practice. I will conclude with recommendations for further research.

Policy Recommendations

A key recommendation in my conclusions is the need for professional development and learning communities. We all need to be supported by each other. My research participants particularly expressed an appreciation of the focus groups in which they could talk with other professionals in their discipline. They need learning opportunities through workshops and conferences. They need support during the school year for guided dialogue and discussion. They need opportunities for mentorship.
“Professional learning communities” is a concept that supports both new and veteran teachers in supporting each other in becoming effective practitioners. Hennessey explains that the components of a professional learning community include a number of practices.

A literature review yields several descriptors of such communities that are identifiable in our culture, including: (1) opportunity to discuss ideas and perspectives collectively; (2) development of shared norms and values; and (3) a willingness to be collectively responsible for initiatives that ultimately lead to excellence in teaching and learning. Additionally, staff development is valued, differentiated opportunities for growth is provided, and there is a norm of continuous growth and improvement. (2005, p. 36).

Collaboration is a key element of professional learning communities that support teachers. Reflective dialogues are also important. The process of discussion and reflection in a group setting can assist a teacher in developing and articulating her thoughts and knowledge about students just as the teachers in Carini’s descriptive child review model were able to clarify information about students in a group setting.

Merton (1955) explains the importance of supporting each other in his classic work which borrows from Donne in stating “No Man is an Island”.

Only when we see ourselves in our true human context, as members of a race which is intended to be one organism and “one body,” will we begin to understand the positive importance not only of the successes but of the failures and accidents in our lives. My successes are not my own. My way to them was prepared by others.... Nothing at all makes sense, unless we admit, with John Donne, that “No man is an island, entire of itself, every man is a piece of the continent, a part of the main.” (Merton, 1955, pp. xxi – xxiii)

Another key recommendation is to provide more time for teachers. They need a slower pace. Assessing students with honesty and respect involves time and reflection (Carini, 2001; Cobb, 2005; Tierney. 2005). All my participants discussed how busy they were. Mueller addresses the negative side of busy jobs and lives.
Our lack of rest and reflection is not just a personal affliction. It colors the way we build and sustain community, it dictates the way we respond to suffering, and it shapes the ways in which we seek peace and healing in the world. I have worked for twenty-five years in the fields of community development, public health, mental health, and criminal justice. With a few notable exceptions, the way problems are solved is frantically, desperately, reactively, and badly. (2000, p. 3).

Mueller explains that the busyness “can become a kind of violence” (p. 5); whereas a slower pace that allows for reflection and the ability to listen can be healing not only to the individual, but those around him or her. Palmer (2007) supports a view of teaching that is reflective and listens from the heart.

The need to include the voice of teachers and minority communities in policy and accountability systems is a more difficult area to address. Accountability systems are political and policy issues. Policy is established through laws, court decisions, and departments of education. Legislation, though, in a democratic society is influenced by all citizens. Plato (1999) did not believe in democracy because of the requirement that all citizens must be educated and informed to make the appropriate decisions to ensure a fair and knowledgeable governing body with appropriate laws.

I would argue that, given the fact that we do live in a democracy, it is all our responsibility to learn and know about current policy and legislation in an effort to advocate for change. Policies are changed and voices are heard by individuals, one at a time, and together, advocating for change. I would also argue that it is our responsibility to advocate for those who do not have a voice. Language minority communities may know how to express their wishes and viewpoints. We can teach them how to express themselves.
We can also advocate for educational legislation and policies that include the voices of minority communities. Legislation can also be drafted that supports the voices of teachers and perspectives that honor teacher knowledge. Legislation can also be drafted that includes assessment and accountability systems that incorporate language learning and assessment theory that affirms the complexity and humanness of language. Accountability can include both large scale assessment and achievement goals, but also other measures of success. Schools can be evaluated on other forms of student performance than those determined from test scores. Teacher developed assessments can be included in the picture. You and I, and the teachers themselves, need to speak out, though, for this to happen.

In a more general sense I believe that by creating an environment that affirms and supports individuals, they will have a voice. We need teachers supported and empowered. We need to talk together and work together. We need time to reflect. We need time to listen to each other.

Within these recommendations for teachers and the educational community, recommendations for working with the communities of ELLs surfaces. Educators that are empowered supported and encouraged to listen, will be encouraged to listen to their students and their families. We also need to encourage the communities to allow for their empowerment and support them in expressing a voice in policy. Teaching and assessing ELLs starts with each one of us. We must create a better system that sees the complete student, with strengths, abilities and potential, as well as their needs.
Research Recommendations

Several research agendas surface from my assertions and recommendations. Several recommendations for research can be developed from the assertions and recommendations that relate to teachers. Data documenting the effectiveness of professional development and learning communities for teachers of ELLs will assist in determining whether the recommendation for support is appropriate. Case studies of ESL and bilingual teachers supported in articulating assessment information would yield valuable data on the benefit of providing support.

Research involving language minority communities that have been involved in the educational policies of the local school district will help inform policies on how to include their voice. Research can also take the form of telling us what students and families want. Research can provide information on what educational goals language communities have for their children.

A recommendation for research that is both more complex and more concrete than the others involves the area of English language proficiency assessment and theoretical foundations. There is definitely a need for more information on how to assess English language proficiency. There is a need for research on what should be included in the large scale assessments, along with research on the most effective informal, authentic assessments.

Along with a recommendation for research on English language proficiency assessment is the need for more research on the construct of English language proficiency. I found different perspectives among my participants on what constitutes proficiency in English for their students. Recent reviews of state assessment systems
have found “substantial variation” (Wolf et al., 2008c, p. 4) in state and school district policies and assessments, causing situations where the same student would be classified at different levels of proficiency in different locations.

While I am hesitant at this point to recommend a uniform definition of proficiency and a single assessment for English language proficiency nationwide, since this may contradict the theoretical basis of my research and recommendations, which allow for multiple perspectives, there is need for more information. More research needs to be conducted on what the construct of English language proficiency involves. This research would include what we should expect for our students and what a view of proficiency that includes multiple voices, including those who advocate for native language preservation and bilingualism looks like.

Summary

The growth of any craft depends on shared practice and honest dialogue among the people who do it. We grow by private trial and error, to be sure – but our willingness to try, and fail, as individuals, is severely limited when we are not supported by a community that encourages such risks. (Palmer, 2007, p. 148)

In summarizing my final chapter of my research study I will use the journey metaphor again. As discussed, this research project was a journey that took me literally and figuratively in wide directions. I started with ideas and questions and built an explorative framework to start my journey. I looked at what the literature reported on my topic. I entered the field, with theory as a guiding light, but open to what was happening in the area I was exploring. I drove highways and country roads to talk with my participants. I gained information which was again reviewed against the literature. I
experienced some failures and also some successes. I developed findings from the data. I ended this part of my journey with assertions and conclusions. I found I have only begun my journey. Some answers to my research questions emerged. In many ways, though, there are more questions. Though there are commonalities in my findings and concepts that can be verified by the literature, there are also areas in which one can interpret multiple meanings. There are contradictions and paradoxes. These questions will lead to further research, exploration and new journeys. The journey of this research study has ended, but I am hopeful for the new journeys that will take place tomorrow, new conversations, and the possibility of new answers and knowledge.
APPENDICES
Appendix A
Consent Form for Participants

Researching English Language Proficiency Assessment Practices

Dear:

You are invited to participate in a research study being conducted by myself, in my role as a student, under the supervision of my advisor, Dr. Shelby Barrentine of the University of North Dakota, College of Education and Human Development, Department of Teaching and Learning. The purpose of the research is to assist me in the completion of requirements for Ph.D. in Teaching and Learning. The information gained from the study could be used to benefit current assessment practices for language minority students. But, the primary purpose in conducting the research is to learn research techniques and complete my program requirements.

The focus of this study is to gather information on the English language proficiency assessment of students. This involves gathering information from you on your practices and ideas on English language proficiency assessment through personal interviews, a questionnaire and focus group conducted as an online chat. The interviews will involve two, or possible three, sessions of 30 to 40 minutes a session. They will be tape recorded. The questionnaire will be sent by mail and will take approximately 15-20 minutes. The focus group will involve the use of a computer with internet capability and will last 60 to 90 minutes. All activities will occur over a period of three to four months.

The questions used in the interview, questionnaire and focus group are not intended to be controversial, personal or in any way stressful. They are related to your thoughts, ideas and opinions. If you are uncomfortable with a question, you have no obligation to answer. Hopefully the experience of participating in this study will be beneficial and positive. Ideally, the experience will lead to a greater understanding of assessment practices for multicultural children our schools.

Participation in this study is completely voluntary. I am not providing financial compensation. You are free to leave the study at any time without penalty. Participation, lack of participation or withdrawal from the study will not influence your employment or
your relationship with any educational institution in the state, including the North Dakota Department of Public Instruction.

Any information from this study that can be identified with you will remain confidential and will be disclosed only with your permission. All data and consent forms will be kept in separate locked cabinets for a minimum of three years after the completion of this study. Only I, the researcher, my advisor, Dr. Barrentine, and people who audit Institutional Review Board (IRB) procedures will have access to the data. After five years, the data will be erased and shredded.

If you have questions about the research, please feel free to call me at 701-221-2572 or Dr. Shelby Barrentine at 701-777-3243. If you have any other questions or concerns, call the Research Development and Compliance office at 777-4279. Please sign below, if you are willing to participate in this study. You will be given a copy of this consent form for future reference.

Sincerely,

Mari B. Rasmussen

All of my questions have been answered and I am encouraged to ask any questions that I may have concerning this study in the future. I understand what the study involves and I agree to participate.

Participant’s signature and date
Appendix B
Research Tools: Questionnaire and Tierney’s Principles of Assessment

ENGLISH LANGUAGE PROFICIENCY ASSESSMENT
A Questionnaire of Teacher Practices and Viewpoints
Winter 2007

Part I: General Questions on English Language Proficiency Assessment Practices

1. Information I use to understand my students’ level of English language proficiency is primarily gathered through formal, large-scale, standardized/standards-based assessment.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

2. Information I use to understand my students’ level of English language proficiency is primarily gathered through authentic or classroom based assessments.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

3. Information I use on my students’ level of English language proficiency is primarily gathered through a combination of large scale and standardized tests and authentic or classroom based assessments.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

4. I do not use assessment information with my students.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>
5. Language proficiency assessment information should be used to modify instruction for students.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Language proficiency assessment information should be used to develop instructional student plans.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. Language proficiency assessment information should to track student growth and progress.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. Language proficiency assessment information should be shared with content area and classroom teachers.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

9. Other purposes for Language proficiency assessment information. (Please explain.)
10. How often do you supplement English Language proficiency test scores with other sorts of assessments to understand what students need and how they are doing?

<table>
<thead>
<tr>
<th>Never</th>
<th>Rarely</th>
<th>Seldom</th>
<th>Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11. Norm-referenced standardized English language proficiency test reports.

<table>
<thead>
<tr>
<th>Never</th>
<th>Rarely</th>
<th>Seldom</th>
<th>Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

12. Academic achievement test information (for example - the language arts portion of the state assessments or other content area testing).

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

13. Norm-referenced standardized reading test (for example - Gates MacGinitie, DIBILS, DRA, etc.)

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14. Authentic or performance assessments (for example - running record, observation, miscue analysis, etc.).

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

15. Other? (Please explain.)
Part II: Indicators of English Language Proficiency

A. What are the best indicators that a student who has come from a non-English background has reached proficiency in English?

16. Student can converse in English adequately.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

17. Student understands what he or she reads.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

18. Student can write in English.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

19. Student can read English texts at a grade level appropriate for his or her age.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

20. Student discusses abstract concepts in English.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>
21. Student quits using his or her home language.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

22. Student scores “advanced” or “proficient” on a norm-referenced English language proficiency test.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

23. Student scores “proficient” in Reading/Language Arts on the state standards–based academic achievement assessment.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

24. Student scores at grade level on a norm-referenced reading test.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

25. Student maintains a “C” or better in academic content areas in the classroom.

<table>
<thead>
<tr>
<th></th>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

26. Other (Please explain.)
Part III: No Child Left Behind (NCLB) and English Language Learners

A. NCLB Assessment Requirements

28. The No Child Left Behind Legislation requires that states and school districts develop new English language proficiency assessments based on state standards and state definitions of proficiency. How knowledgeable about your state’s English language proficiency assessment program are you?

<table>
<thead>
<tr>
<th>Not at all Knowledgeable</th>
<th>Not Very Knowledgeable</th>
<th>Neutral</th>
<th>Somewhat Knowledgeable</th>
<th>Very Knowledgeable</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

Please Explain:

29. The No Child Left Behind Legislation requires that states involve teachers, administrators and parents in the development process of state assessments. Have you been involved with this development process?

<table>
<thead>
<tr>
<th>Not at all Involved</th>
<th>Not Very Involved</th>
<th>Neutral</th>
<th>Somewhat Involved</th>
<th>Very Involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

Please Explain:

30. The No Child Left Behind Legislation requires that states and school districts disaggregate the data for academic achievement on students who are limited in English language proficiency achievement goals. Is it useful for you to see how the limited English proficient students do in comparison to other students?

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

Please Explain:
31. The No Child Left Behind Legislation requires that states report on progress students made toward English language proficiency objectives based on the state English language proficiency assessment. Is this information helpful to you?

Not at all Useful | Not Very Useful | Neutral | Somewhat Useful | Very Useful
---|---|---|---|---
□ | □ | □ | □ | □

Please Explain:

B. What information would you like to see from NCLB state English language proficiency assessments?

32. Description of student’s overall level of English language proficiency.

Not at all Important | Not Very Important | Important | Somewhat Important | Very Important
---|---|---|---|---
□ | □ | □ | □ | □

33. Information on how the student performed in the language sub-domains of reading, writing, speaking, and listening.

Not at all Important | Not Very Important | Important | Somewhat Important | Very Important
---|---|---|---|---
□ | □ | □ | □ | □

34. Information on how the student performed against English Language Proficiency Standards.

Not at all Important | Not Very Important | Important | Somewhat Important | Very Important
---|---|---|---|---
□ | □ | □ | □ | □
35. Student growth in language proficiency from previous year.

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

36. Comparison with other students limited in English in the school district.

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

37. Comparison with other students limited in English in the state.

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

38. Recommendations for instruction.

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

38. Other? (Please explain.)
Part IV: Participant Demographics

1. What position do you currently hold?
   □ ESL Teacher
   □ Bilingual Education Teacher
   □ Other -

2. What grade level(s) do you work at?
   □ K-3
   □ 4-6
   □ 6-8
   □ 9-12
   □ K-12

3. Years of experience in education:
   □ 1 year
   □ 2-5 years
   □ 5-10 years
   □ 10-15 years
   □ 15-20 years
   □ 20-25 years
   □ 25-30 years
   □ 30+ years

1. Educational level
   □ Bachelor’s Degree
   □ Master’s Degree
   □ Specialist Degree
   □ Doctorate Degree

2. School District Enrollment. The school district I work in has an enrollment of:
   □ 1-1000 students
   □ 1000-2500 students
   □ 2500-5000 students
   □ 5000-7500 students
   □ 7500-10,000 students
   □ 10,000-12,500 students
   □ 12,500 or more

3. The ethnic backgrounds of the students who are English Language Learners in the school where I work are primarily:
   □ Wide diversity of ethnic groups
   □ Hispanic
   □ German
   □ Native American
   □ Other

THANK YOU FOR YOUR TIME FILLING OUT THIS QUESTIONNAIRE!!!
TIERNEY'S PRINCIPLES OF ASSESSMENT

Principle 1: Assessments should emerge from the classroom rather than be imposed upon it.

Principle 2: Effective testing requires teacher professionalism with teachers as learners.

Principle 3: Assessment practices should be client centered and reciprocal.

Principle 4: Assessment should be done judiciously, with teachers as advocates for students and ensuring their due process.

Principle 5: Assessment extends beyond improving our tests to the purposes of assessment and how results from assessment are used, reported, contextualized and perceived.

Principle 6: Diversity should be embraced, not slighted.

Principle 7: Assessment procedures may need to be non-standardized to be fair to the individual.

Principle 8: Simple-minded summaries, scores, and comparisons should be displaced with approaches that acknowledge the complex and idiosyncratic nature of literacy development. Straightforward comparisons across individuals are usually arbitrary, biased, and narrow.

Principle 9: Some things that can be assessed reliably across raters are not worth assessing; some things that are worth assessing may be difficult to assess reliably except by the same rater.

Principle 10: Assessment should be more developmental and sustained than piecemeal and shortsighted.
Principle 11: Most interpretations of results are not straightforward. Assessment should be viewed as ongoing and suggestive, rather than fixed or definitive.

Principle 12: Learning possibilities should be negotiated with the students and stakeholders rather than imposed via standards and assessment that are preset, prescribed, or mandated.

Principle 13: Assessment should be assessed in terms of its relationship with teaching and learning, including the opportunities learners are offered and the rights and respect they are accorded.

(Tierney, R. J., 2005, PP. 23 – 40)
Appendix C
Research Protocols: Interview and Focus Group Protocol
And
Online Focus Group Protocol

INTERVIEW AND FOCUS GROUP PROTOCOL

"The research interview is based on the conversations of daily life and is a professional conversation. One form of research interview—a semistructured life world interview— is defined as an interview whose purpose is to obtain descriptions of the life world of the interviewee with respect to interpreting the meaning of the described phenomena" (Kvale, 1996, pp. 5–6).

Consistent with Kvale's description of conversation as research, the protocol I followed in interviewing will involve typical conversational practices.

First Interview

The first interview involved general discussion of English language proficiency assessment. It will also involve discussion of the interviewee's background and experiences with language minority students and assessment. I will have summarized data from the questionnaire and we will discuss the results. Seidman explains that "In the first interview, the interviewer's task is to put the participant's experience in context by asking him or her to tell as much as possible about him or herself in light of the topic up to the present time" (1998, p. 11). Questions included:

- Tell me about your teaching situation.
  - What kind of students have you taught?
  - How have things changed in your teaching career?
- What have you used to assess the language proficiency and progress of the students with whom you teach?
  - What works for you?
• What doesn’t work?

• What does English language proficiency mean to you?

• What sort of formal tests do you use?
  
  o What information do these tests provide?
  
  o Is it helpful?

• What sort of authentic or alternative assessments have you used?

• Based on the questionnaire conducted, most of the teachers participating in the survey believe ____________ about English language proficiency assessment. What are your thoughts?

Second Interviews

The second interviews followed up with some of the themes that surfaced in the first interview and focus group and involve more in depth questions. Specific questions are included in Table 5 in Chapter III.

Online Focus Groups

The online focus groups were conducted as chats using the live chat feature of a Course Management System, available through the University of North Dakota; Blackboard. The communication sent out prior to the scheduled focus groups follows.
ONLINE FOCUS GROUPS
Research Project for Mari Rasmussen
University of North Dakota
Tuesday, June 12th

Group #2
2:30 – 3:30pm Central Daylight Time (1:30 – 2:30pm Mountain Daylight Time)
Group #3
4:00 – 5:00pm Central Daylight Time (3:00 – 4:00pm Mountain Daylight Time)

Purpose of the Online Focus Group:
Discuss with other teachers through a virtual focus group environment issues related to English language proficiency assessment.

Directions to participate in the Online Focus Group:
1. Go to the website: www.online.und.edu
2. Go to “user login” on the left side of the screen
3. Log on with your name: firstnamelast name (no dots or spaces)
4. Use the Password “elpa”
5. Once you login, it is recommended that you change passwords.
   Under “My Courses”, go to “English Language Proficiency Assessment: Teachers Perspectives”.
6. Go to “Online Focus Group Chat” in the green area on the left side.
   Documents, including Directions & Questions, Tierney’s Principles of Assessment and Questionnaire Results are also available on the left side.
7. Click “join” on the screen to join the group.
8. Discussion involves typing questions and comments.
   If you have problems, call Chad Bushy at 701-777-2728 or send an e-mail message to chadbushy@mail.und.nodak.edu.
Items to discuss:

Questionnaire Results: A majority of teachers use authentic or classroom-based assessments to learn about the English language proficiency of the students.

Question:
• What do these assessments reveal about your students?

A majority of teachers use a combination of large scale and standardized tests, along with authentic assessments to learn about students and their language proficiency. Many saw the importance of using both. Some teachers have worked with states in developing assessments and setting policy. But, not all teachers participate in state activities.

Question:
• What is the relationship between teacher developed assessment information and large scale assessment and accountability?

Responses differed in a couple areas on the questionnaire. Ideas varied on Indicators of English language proficiency.

Question:
• What characteristics indicate proficiency to you?

Responses varied on the use of the large scale state tests required by No Child Left Behind Act.

Questions:
• Is it helpful to compare English Language Learners with other English Language Learners?
• Is it helpful to compare English Language Learners with non-English Language Learners?

Tierney’s Principles of Assessment

Choose one of Tierney’s principles of assessment to discuss
Appendix D
Questionnaire Results

ENGLISH LANGUAGE PROFICIENCY ASSESSMENT
A Questionnaire of Teacher Practices and Viewpoints
RESPONSES

Part I: General Questions on English Language Proficiency Assessment Practices

A. Sources of Information

5. Information I use to understand my students' level of English language proficiency is primarily gathered through formal, large-scale, standardized/standards-based assessment. N = 11

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

6. Information I use to understand my students' level of English language proficiency is primarily gathered through authentic or classroom based assessments. N = 11

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>

7. Information I use on my students' level of English language proficiency is primarily gathered through a combination of large scale and standardized tests and authentic or classroom based assessments. N=11

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

261
8. I do not use assessment information with my students
N = 11

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

B. Use of English Language Proficiency Assessment Information

9. Language proficiency assessment information should be used to modify instruction for students
N = 11

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>7</td>
</tr>
</tbody>
</table>

10. Language proficiency assessment information should be used to develop instructional student plans
N = 11

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

11. Language proficiency assessment information should to track student growth and progress
N = 11

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

262
12. Language proficiency assessment information should be shared with content area and classroom teachers

N = 11

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

9. Other purposes for Language proficiency assessment information. (Please explain.)
- Identification purposes
- Provides statistics that can be used to compare LEP students to their LEP peers and mono-lingual peers.
- Ability to problem-solve – social interactions, settings.
- Fulfill state requirements – assess level for grouping
- Evaluate district program
- Educate mainstream, pre-service teachers and public regarding the challenges of ELLs
- To evaluate instructional materials or reading, literacy, phonics curriculum; to provide information about language acquisition.

10. How often do you supplement English Language proficiency test scores with other sorts of assessments to understand what students need and how they are doing?

N = 11

<table>
<thead>
<tr>
<th>Never</th>
<th>Rarely</th>
<th>Seldom</th>
<th>Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

- We just did our first test and have not results
- 2006 was the first year Montana implemented a language proficiency assessment. The results have not been released.

C. Types of assessment information
What assessment information on your students is most helpful for you?

11. Norm-referenced standardized English language proficiency test reports

N = 11

<table>
<thead>
<tr>
<th>Never</th>
<th>Rarely</th>
<th>Seldom</th>
<th>Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

- Not yet
- IPT
12. Academic achievement test information (for example - the language arts portion of the state assessments or other content area testing)  
N = 11

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

13. Norm-referenced standardized reading test (for example - Gates MacGinitie, DIBELS, DRA, DRP, etc.)  
N = 11

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>

14. Authentic or performance assessments (for example - running record, observation, miscue analysis, etc.)  
N = 11

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

15. Other? (Please explain.)  
No responses
Part II: Indicators of English Language Proficiency

A. What are the best indicators that a student who has come from a non-English background has reached proficiency in English?

16. Student can converse in English adequately
N = 11

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

17. Student understands what he or she reads
N = 11

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

18. Student can write in English
N = 11

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

19. Student can read English texts at a grade level appropriate for his or her age
N = 11

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>
20. Student discusses abstract concepts in English  
\(N = 11\)

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>9</td>
</tr>
</tbody>
</table>

21. Student quits using his or her home language  
\(N = 11\)

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

22. Student scores “advanced” or “proficient” on a norm-referenced English language proficiency test  
\(N = 11\)

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

23. Student scores “proficient” in Reading/Language Arts on the state standards–based academic achievement assessment  
\(N = 11\)

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

24. Student scores at grade level on a norm-referenced reading test  
\(N = 11\)

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>
25. Student maintains a "C" or better in academic content areas in the classroom  
N = 11

<table>
<thead>
<tr>
<th>Does not at all Indicate Proficiency</th>
<th>Not a Very Good Indicator of Proficiency</th>
<th>May be an Indicator of Proficiency</th>
<th>Somewhat an Indicator of Proficiency</th>
<th>Strong Indicator of Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

26. Other (Please explain.)  
No Responses

Part III: No Child Left Behind (NCLB) and English Language Learners

A. NCLB Assessment Requirements

39. The No Child Left Behind Legislation requires that states and school districts develop new English language proficiency assessments based on state standards and state definitions of proficiency. How knowledgeable about your state’s English language proficiency assessment program are you?  
N = 11

<table>
<thead>
<tr>
<th>Not at all Knowledgeable</th>
<th>Not Very Knowledgeable</th>
<th>Neutral</th>
<th>Somewhat Knowledgeable</th>
<th>Very Knowledgeable</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Please Explain
- TEAE
- ND State ELL Advisory Committee
- I know we do the TEAE test, but am not sure if it is a state, national or district test. We also do the SOLOM.

40. The No Child Left Behind Legislation requires that states involve teachers, administrators and parents in the development process of state assessments. Have you been involved with this development process?  
N = 11

<table>
<thead>
<tr>
<th>Not at all Involved</th>
<th>Not Very Involved</th>
<th>Neutral</th>
<th>Somewhat Involved</th>
<th>Very Involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Please Explain
- I have not participated in item analysis or setting cut scores, but am the district test coordinator.
• Served on western state consortium test development committee; trained to administer WIDA ACCESS; served on alignment committee; served on advisory board

41. The No Child Left Behind Legislation requires that states and school districts disaggregate the data for academic achievement on students who are limited in English language proficiency achievement goals. Is it useful for you to see how the limited English proficient students do in comparison to other students?

N = 11

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Please Explain
• In writing effective reports and setting school goals.
• Yes, I can forget what “average” kids are capable of.

42. The No Child Left Behind Legislation requires that states report on progress students made toward English language proficiency objectives based on the state English language proficiency assessment. Is this information helpful to you?

N = 11

<table>
<thead>
<tr>
<th>Not at all Useful</th>
<th>Not Very Useful</th>
<th>Neutral</th>
<th>Somewhat Useful</th>
<th>Very Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Please Explain:
• Yes, now our kids count and can’t be ignored.

B. What information would you like to see from NCLB state English language proficiency assessments?
• Recommendations for what to do with the information.

43. Description of student’s overall level of English language proficiency

N = 11

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
44. Information on how the student performed in the language sub-domains of reading, writing, speaking, and listening
N = 11

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

45. Information on how the student performed against English Language Proficiency Standards.
N = 11

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

46. Student growth in language proficiency from previous year.
N = 11

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

47. Comparison with other students limited in English in the school district.
N = 11

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

* Problems may indicate other areas of concern – SPED, emotional, family

48. Comparison with other students limited in English in the state
N = 11

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>
49. Recommendations for instruction
N = 11

<table>
<thead>
<tr>
<th>Not at all Important</th>
<th>Not Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Very Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

- What would this look like? Too general – no help or standards that need to be addressed and suggestions for how to do what is good.

38. Other? (Please explain.)
- We haven’t received any data on our first assessment.
REFERENCES


276


277


278


King, L. (2008, June 25). Test scores are up, but is it No Child? *USA Today*, p. 7D.


283


287


