Making sense of testing: English language learners and statewide assessment

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE
UNIVERSITY OF MINNESOTA
BY

Karla Rae Stone

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Martha Bigelow, Adviser

July 2008

ACKNOWLEDGEMENTS

i

I have extraordinary gratitude for the encouragement and role modeling set by Tina Edstam –whose passion for sharing her knowledge of working with English language learners is an example to which I aspire.

Finally, I am exceptionally grateful for Bill and Ethan, without whom none of this would have been possible.

DEDICATION

To Ethan - whom I will always love more than anything.

Karla Rae Stone   Word Count 346

ABSTRACT

The federal No Child Left Behind (NCLB) Act of 2001 has dramatically changed the educational landscape for all students by increasing the number of standardized tests used for accountability purposes. The impact is profound on students for whom English is not their first language, in part because of the sheer number of tests they are required to take. The challenges increase when tests become high stakes for students in the form of high school exit exams, an increasingly prevalent requirement in states, though not required under NCLB.

The policies for tests are complex and can easily be viewed as devoid of any attention to the human side of testing – the ground level, nuanced, and sometimes complicated application of policy at the student level. Yet it is at this level where a policy becomes truly validated. The literature on accountability testing rarely considers student perceptions. This research is an attempt to contribute to that void by combining a student voice paradigm with a sense-making framework.

Case study methodology was used and incorporated methods that moved from breadth towards depth as the study evolved. In the case of this research, that evolution was from participant observation (Phase I) to document review, inquiry groups and interviews (Phase II).

Participants' pre-existing knowledge and experiences with testing are explored alongside the new knowledge gained as participants in this research study.

Findings clearly demonstrate that the stakes connected to testing influence the ways in which students make sense of testing, as do the policy signals.

Beyond contributing to the fields of educational policy and student voice initiatives, the intent of this research is to spur advocacy on behalf of and in conjunction with English language learners. Whether for systems accountability (the Test of Emerging Academic English), or for individual accountability as a requirement for a diploma in Minnesota (the Basic Skills Tests), the findings reveal both the complexities of and urgency for exploring testing policies from students points of view. Additionally, findings validate the need for students' voices in policy research, and suggest several implications based on reported experiences with statewide testing.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

IINTRODUCTION

What would happen if we treated the student as someone whose opinion mattered in the introduction and implementation of reform in schools?
(Fullan, M. 1991, p. 170)

Background

Students are the intended recipients and key stakeholders of large-scale testing policy. Yet they tend to be viewed as passive recipients of reform. As Michael Fullan suggests in the opening quote, and others reiterate, students have opinions about the policies to which they are often subjected (Mitra, 2007; Rudduck, 2007; Thiessen, 2007; Yonezawa & Jones, 2007). This study is predicated on the belief that students ought to be included as active participants in the implementation and examination of testing policies, and not simply thought of as passive beneficiaries of them. Students ultimately play a critical role in the success or failure of any educational policy's implementation, and are deeply affected by both intended and unintended consequences of testing policy. Seldom explored aspects of policy implementation are the inherent assumptions of basic understanding, minimal engagement and genuine cooperation on the part of the students.

There have been recent and repeated calls for the inclusion of students' voices in educational research, generally, and in large-scale school reform initiatives (Cook-Sather, 2002, 2007; Fielding, 2001; Lee & Simon-Maeda, 2006; Oldfather, 1995; Rubin, 2003; Rudduck, 2007; Silva, 2003; Thiessen, 2007; Yonezawa & Jones, 2002, 2007). The same

1

call for their inclusion is not as evident in the examination of large-scale testing policy, or policy implementation research generally. Students are key stakeholders of such policy and bring a host of individual characteristics, opinions and understandings to policy implementation. Multidimensional factors such as low achievement, effort and motivation, (complicated by first language and literacy in the case of English language learners), may impact students' engagement with tests. So often testing policies put the onus for achievement on students and "ultimately ignore the complexities of children's lives" (Roderick & Engel, 2001, p. 221).   Ignorance of such complexities results in blind faith in the ultimate validity of results. So how is one to uncover the complexities? As Graue and Walsh (1998) note, "the American cultural obsession with being practical works against finding it out…truly finding out requires researchers to look in avoided places and in unfamiliar ways… Why study children? To find out" (p. xv). Given the paramount place of large-scale testing in US schools today, the time has come to explore the experiences of students with tests from their own perspectives.

If one views policy as iterative and open to re-examination (or in the case of No Child Left Behind - reauthorization), then allowing students an entrée into the process may illuminate inconsistencies in policy design, implementation and the resulting consequences. Engaging students in dialogue about testing policy can provide opportunities for students to reflect on and make sense of their experiences collectively, building mutual understandings of a policy's intent. Research including student

perspectives provides both a methodological and epistemological orientation for an examination of policy from students' points of view.

Student voice research encompasses studies which seek students' perceptions, perspectives, and opinions at any point along a continuum of student engagement. A student voice research paradigm recognizes that while students are "the recipients of reformers' good intentions, [they] are infrequently asked their opinions as to what enhances or detracts from their learning" (Yonezawa & Jones, 2007, p. 681). Dennis Thiessen describes how student voice research allows participants to "critically examine the individual and collective worlds of students, with a particular focus on how they navigate and negotiate the dynamic and multi-dimensional demands of their classroom and school lives" (2007, p. 5). Additionally, qualitative studies with adolescent English language learners – as one subset of the student population - can elicit the often silenced, marginalized voices of those whose agency is integral to the relative success of testing. The focus of inquiry in this study is statewide accountability testing as one component of those "dynamic and multi-dimensional demands" of students' lives in school.

Practicality and efficiency are epitomized by standardized testing in the US, policies for which are inherently complex. An examination of standardized testing policies from the perspective of students, whose lives are also unique and multi-faceted, requires an attention to students as influential contributors to test policy implementation. Policy implementation relies on the assumption that the policy will be enacted according to its intent. Along the way, however, a number of agents (administrator, teachers, and

3

students, among others) have a great deal of influence and control over the actual implementation of policy and the validity of results. Like educational research, educational policy can be viewed as a complex social practice (Koyama, 2004; Sutton & Levinson, 2001), an "on-going process of normative cultural production constituted by diverse actors across diverse social and institutional contexts" (Sutton & Levinson, 2001, p.1). Policy implementation research has explored the sense teachers make of policy, but has stopped short of including students as agents in that process (Abernathy, 2007; Louis, Febey & Schroeder, 2005; Schmidt & Datnow, 2005; Spillane, Reiser & Reimer, 2002). Additionally, most previous studies of statewide testing and students do not consider English language learners as key participants (Chabran, 2003; Flores & Clark, 2003; Wheelock, Bebel & Haney, 2000). Of the studies that do, the active role of these learners in the testing process has not been considered from the students' points of view (Anderson, 2004; Magnuson, 2003; Menken, 2006; Monroe, 2006). Overall, little has been done in the way of research exploring the opinions of those who stand to gain or lose the most by recent educational policy and the resulting emphasis on accountability tests: the students.

Within a student voice research paradigm, a sense-making framework borrowed from policy implementation studies is utilized for this study. It explores the multi-faceted variables that impact the ways in which English language learners come to understand statewide testing.

Overview of Study

Anderson (2004) notes how qualitative research allows for investigation into "how people make sense of an aspect of the world" (p. 46). It is the intent of this research to turn the investigative lens towards adolescent English language learners (ELL), and the extent to which they come to understand and therefore engage with statewide testing policy. This case study explores the ways in which high-school ELL make sense of statewide testing based on the fundamental knowledge they have of the tests, their experiences with the process, and the understanding they develop through a process of dialogic reflection and inquiry. The following research questions are explored in this dissertation:

- What *knowledge* do high-school aged English language learners (ELL) in

  a suburban Minnesota school district have of statewide assessments?

- How do high-school aged ELL in a suburban Minnesota school district

  describe their *experiences* with statewide assessments?

- How do high-school aged ELL *make sense* of statewide assessments?

Grounded in a student voice orientation to research, methods were used that can be described as 'funnel shaped' in coverage (McCall-Perez, 1991), meaning the focus moved from breadth towards depth as the study evolved. In the case of this research, that evolution was from participant observation (Phase I) to document review, inquiry groups and interviews (Phase II). The primary source of data emerged during the methods which

5

sought depth, namely inquiry groups and interviews. Three separate groups of English language learners (11 participants in total) comprised the inquiry groups, and each group met three times during the spring of 2007. These groups provided an opportunity for reflection and co-construction of understandings around the topic of statewide testing. Data from inquiry groups is supported by that from other data sources in the report and the discussion of findings.

Because of ever changing requirements and the sheer number of tests taken by English language learners, a deliberate attempt was made to focus the current study. To that end, tests that are the result of either federal and/or state testing policies are explored with students. Background information and highlighted sections of relevant testing policies, both federal and state, are described in the next section.

United States Testing Policy and English Language Learners

The federal No Child Left Behind (NCLB) Act of 2001 has dramatically changed the educational landscape for all students by requiring increasing the number of standardized tests used for accountability purposes. The impact is profound on students for whom English is not their first language, in part because of the sheer number of tests they are required to take**.** The challenges increase when tests become high stakes for students in the form of high school exit exams, an increasingly prevalent requirement in states, though not required under NCLB. In fact, Minnesota is one of twenty-two states requiring passing scores on exit exams for graduation (Center for Education Policy,

2006). Significant gaps have been reported nationwide between the pass rates of English language learners and the overall pass rate (Short & Fitzsimmons, 2007). These tests can be an additional burden if requisite attention is not paid to the linguistic needs of students (LaCelle-Peterson, 2000), thereby making the challenge of passing the tests even greater for English language learners. The stakes are extraordinarily high for English language learners in high school.

While others have described all testing under federal education policy as high stakes (Menken, 2006; Nichols & Berliner, 2007), this research makes a distinction between those tests that have particularly serious consequences for students, and those for which the stakes are higher for schools. Magnuson (2003) refers to this distinction as tests for *systems* accountability versus *individual* accountability. A discussion of federal and Minnesota-specific testing policies follows with the aim to further elucidate the distinction.

Accountability for the academic success of all students, including English language learners, became an integral part of the 1994 reauthorization of the Elementary and Secondary Education Act (ESEA), which was known as the Improving America's Schools Act (IASA). The focus on accountability became even more stringent with the 2001 reauthorization –NCLB - (Rivera & Collum, 2006) when accountability became closely intertwined with sanctions. The 2001 law, among other things, provides a much greater emphasis on English language proficiency testing, requiring that states "provide for annual assessment of English proficiency (measuring students' oral language, reading,

7

and writing skills in English) of all students with limited English proficiency in the schools served by the State educational agency" (sec. 1111 [b][7]). In another change from IASA, Title VII was replaced with Title III, and the resulting exclusive emphasis on English proficiency figures prominently in the requirements for accountability (Menken, 2006; Wright, 2005).

No Child Left Behind has continued the push that began with IASA for all states to create challenging academic content and student academic achievement standards (Table 1.1), along with increasingly prescriptive requirements regarding the assessments that measure accountability for those standards. Data from state assessments must now be disaggregated, bringing a new focus to various subgroups of students, including ELL[1] that was lacking in previous iterations of ESEA. Previous requirements, lacking both a focus on vulnerable students and significant consequences, meant that the low achievement of ELLs and other under-performing subgroups was often masked by aggregated data.

---

[1] In the context of writing about and citing federal and state legislation, the term limited English proficient (LEP) will be used. Elsewhere in this dissertation, the more common terminology of English language learners (ELLs) is used.

Table 1.1: Standards, Assessments and Annual Measurable Objective Requirements in NCLB

POPULATION	STANDARDS	ASSESSMENTS	ANNUAL MEASURABLE OBJECTIVES
For LEP Students Only	English Language Proficiency Standards	English Language Proficiency Assessments	English Language Proficiency Annual Measurable Objectives
For ALL Students (including LEP Students)	Academic Content Standards (reading/language arts, mathematics, science) Academic Achievement Standards (reading/language arts, mathematics, science)	Academic Assessments in reading/language arts, mathematics, and science	Annual Measurable Objectives in reading/language arts and mathematics

(U.S. Department of Education, 2003)

The sections of NCLB that have specific relevance to the topic of adolescent English language learners include Title IX Part A (General Provisions: Definitions), Title I (Improving the Academic Achievement of the Disadvantaged), and Title III (Language Instruction for Limited English Proficient and Immigrant Students). Each of these sections will be discussed in light of testing requirements, and help portray the landscape of statewide testing encompassing this study.

The ubiquitous nature of testing under NCLB is such that it is tied to the very definition of LEP students in both federal and state (Minnesota) law. Under Title IX, the definition of a limited English proficient individual is one who:

> (A) is between the ages of 3 through 21, (B) is enrolled in an elementary or secondary school, (C) was not born in the United States or whose native language is a language other than English, and (D) whose difficulties in speaking, reading, writing, or understanding the English language may be sufficient to deny the individual – (i) the ability to meet the State's proficient level of achievement on State assessments described in section 111(b)(3). (Sec. 9101(25)(A-D)).

The definition goes on to relate language difficulties to those that impact both classroom achievement and opportunities for full participation in society.

According to the federal government, this definition gives states "flexibility in defining who constitute the LEP subgroup" (United States Department of Education [DOE], 2004, p. 1). States may choose to include only those students receiving direct, daily service, or expand inclusion to those students who are being monitored based on assessment data. In addition, in 2004 Secretary of Education Paige announced "new flexibility [that] would, for AYP calculations, allow states for up to two years to include in the LEP subgroup students who have attained English proficiency", giving states "the flexibility to allow schools and local education agencies (LEAs) to get credit for improving English language proficiency from year to year" (DOE, 2004, p. 1). In

addition to claims of flexibility, Title III includes a call for research to find a "common definition of limited English proficient child[ren] for purposes of national data collection…" (3222(b)(3)). Lack of a common definition aside, all states are held to the same rhetorical requirements of accountability, though states have discretion in their choice of assessments and in the setting of target scores.

Rivera and Collum (2006) have examined requirements for including ELLs in assessment systems, from the standpoints of inclusion and accountability. They define inclusion requirements as those "designed to ensure that ELLs participate fully and appropriately in state assessments, [and] cover the use of English language proficiency measures, as well as accommodations. Accountability requirements are those that [regard] how ELLs' scores should be reported and to whom these scores should be made available" (p. xxxviii). According to this framework, requirements that fall under inclusion can be found under Title I of NCLB.

Inclusion requirements

Title I (Improving the Academic Achievement of the Disadvantaged), is intended "to ensure that all children have a fair, equal, and significant opportunity to obtain a high-quality education and reach, at a minimum, proficiency on challenging state academic achievement standards and state academic assessments" (Sec. 1001). It is the largest federal program supporting elementary and secondary education (US Department of Education, 2002), and requires annual assessments of all students in reading or language arts and math in grades 3-8, and by 2007-2008, in science "not less than once" during the

11

grade spans 3-5, 6-9, and 10-12 (Sec. 111, [3][C][v][I-II]). While students in grades 3-12 are included in testing requirements, many districts focus Title I funds "on students in the early grades; three-quarters (77 percent) of Title I participants are in preschool through grade 6" (DOE, 2002). Therefore, the policies, especially those regarding testing, apply to students across grade levels, yet funding is often focused primarily at the elementary levels, perhaps to the detriment of adolescent ELLs (especially those who are new to the U.S). In other words, funding is often concentrated at the elementary level, yet academic pressure is greatest for students who are high school – aged. This is especially true for adolescent ELL, who have the least amount of time in which to make significant gains in language proficiency and academic content. According to Ruiz-de-Velasco and Fix, "foreign-born immigrant children now represent a substantially larger share of the total high school student population (5.7 percent) than they do of the primary school population (3.5 percent)" and many of the recently arrived students pose a variety of challenges for schools (2002, p. 20). Among those challenges are varying amounts of prior schooling, varying levels of English language proficiency, and the relatively short timeframe in which adolescent English language learners must face high stakes tests.

Prior to 2002, states could decide when and how to include limited English proficient students in statewide testing, and many different forms of exemptions were allowed for such learners across states' assessment systems. Title I of NCLB requires all LEP students to take the state content tests regardless of language proficiency. However, in 2004 additional flexibility was added, allowing LEP students who have been enrolled

in the US less than one year the option of taking the reading/language arts content

assessment in addition to taking the English language proficiency assessment. Their

scores on the math and the reading/language arts test (if given) would not need to be

included in AYP calculations. Yet "not less than 95 percent of each group of

students…enrolled in school are required to take the [state] assessments" (Sec. 111

[b][3][I][ii]). As stated earlier, states may also include the scores of students who have

attained proficiency for up to two years, offering a glimmer of understanding of the

continuum along which language proficiency lies, yet for reporting purposes the LEP

subgroup still assumes a degree of homogeneity that inaccurately simplifies the

complexity of language learning and language learners (Abedi, Hofstetter & Lord, 2004;

La-Celle Peterson, 2000).

Additional flexibility is allowed for native language testing for up to three years

(or five on a case-by case basis). The reality, though, is that these assessments are often

inappropriate for students who lack a strong literacy background in their first language, or

when appropriate, they are often unavailable. Where native language assessment does

exist, alignment with state standards is not guaranteed, raising a host of validity and

reliability issues. Or, as is the case in Minnesota, cost prohibits state-level translated

versions and the option is passed to districts where finances are equally scarce.

Inclusion in Title I testing requirements has been described as one of the only, if

not the sole positive outcome of NCLB for English language learners (Center on

Educational Policy, 2006; Coltrane, 2002; Gottlieb, 2003, 2006; Lazarin, 2006; Taylor,

Shephard, Kinner & Rosenthal, 2003; Wright, 2005). Though Title I tests are not a focus of this study, it is critical to understand this important contextual piece of the testing puzzle. Findings of this study have implications to Title I tests, as well as for those under Title III and state law.

Accountability Requirements

Federal legislative *accountability* requirements tied to statewide assessments can be found in both Titles I and III. The Title III accountability requirement regarding proficiency and English language learners is a direct focus of this study. Each state must report disaggregated student achievement data to the federal government as well as to the public for reading/language arts and math by subgroup, including limited English proficiency, unless the number of students in a category is "insufficient to yield statistically reliable information or the results would reveal personally identifiable information about an individual student" (Sec. 1111[a][2][C][v][II]). States must also demonstrate that LEP students are making progress towards English language proficiency (Sec. 3121 [a][3]). Furthermore, data must be reported for individual students to "allow parents, teachers, and principals to understand and address the specific academic needs of students", and that report should be "in an understandable and uniform format, and to the extent practicable, in a language that parents can understand" (Sec. 1111[b][3][C][ix-xii]). The noticeable absence of students on the list of stakeholders to whom schools must report data highlights their position as peripheral in terms of agency in the law.

14

Minnesota Testing Requirements and English Language Learners

The Test of Emerging Academic English (TEAE) is a test of reading and writing given to all students (grades 3-12) who are designated limited English proficient in Minnesota. It is used to meet federal requirements for demonstrating progress in language proficiency (Title III) and to fulfill state requirements for identification of and funding for LEP students. The Minnesota Student Oral Language Observation Matrix (MN SOLOM) fulfills the same obligations for tracking students' listening and speaking proficiency.

In Minnesota, in addition to speaking another language at home, a pupil of limited English proficiency is one who…"scored below the state cutoff score on an assessment measuring emerging academic English provided…during the previous school year" and is "determined by developmentally appropriate measures…to lack the necessary English skills to participate fully in classes taught in English" (Chapter 124D.59). Those developmentally appropriate measures might include observations, teacher judgment, parent recommendations or other assessments.

In Minnesota, statewide accountability testing began in 1996 in order to comply with IASA and state law (Anderson, 2004). In addition to meeting Title I accountability requirements, students must pass tests of reading, writing and math in order to receive a diploma. Current requirements for the inclusion of ELL in statewide testing can be found in Chapter 120B.3 of Minnesota Statues 2005. For students enrolled in grade eight before 2005-2006, the Minnesota Basic skills Tests (BST) of reading, writing and math must be passed before graduation, thereby making these tests high-stakes for students. If a student

has been in the U.S. fewer than three years, his or her scores are exempt from reporting (Subd. 1[d]). Under the Basic Skills Testing requirements, for adolescent English language learners time in the U.S. is the only possible exemption from high-stakes testing when the stakes are a high school diploma.

For students enrolled in grade eight in the 2005-2006 year and later, "only the Minnesota Comprehensive Assessments Second Edition (MCA-IIs) in reading, mathematics, and writing shall fulfill students' academic requirements" (Chapter 120B.3 Subd. 1[a][2]). In line with federal testing requirements for accountability purposes, Minnesota statute states the commissioner must require: "(1) annual reading and mathematics assessments in grades 3 through 8 and at the high school level for the 2005-2006 school year and later; and (2) annual science assessments in one grade in the grades 3 through 5 span, the grades 6 through 9 span, and a life sciences assessment in the grades 10 through 12 span for the 2007-2008 school year and later" (Subd. 1a).

Beginning with ninth graders in 2006-07, the Basic Standards Tests are no longer required, but rather the GRAD (Graduation-Required Assessments for Diploma) tests will be. Ninth graders, including ELLs, will take the Writing GRAD. The Reading GRAD will be administered in 10th grade, and combined with the MCA-II beginning in 2007-08. The Mathematic GRAD, also combined with the MCA-II, will be given in 11th grade beginning in 2008-09. In short, dual testing requirements have been collapsed, merging federally required accountability assessments with the state required graduation test into one set of tests which will now assume a duality of purpose.

16

Summary

As demonstrated above, testing policy is complex and can easily be viewed as devoid of any attention to the human side of testing – the ground level, nuanced, and sometimes complicated application of policy at the student level. Yet it is at this level where a policy becomes truly validated. As Rubin and Silva (2003) explain, "without attention to the student experience, we run the risk of reproducing policies and practices that ignore the social character of schooling and undermine the role of students as partners in shaping and changing their own educations" (p. 2). This research is an attempt to help fill that void.

CHAPTER 2

REVIEW OF THE LITERATURE

A focus on outcomes in the United States is certainly more in keeping with the implementation of NCLB, for which the criterion for success is particular scores on standardized tests, where not a student voice is heard.
(Cook-Sather, A. 2006, p. 373)

This review of the literature is intended to frame a discussion of students as both stakeholders and agents in the testing process. It begins with a general examination of the rationale for standardized testing in U.S. public schools coupled with a theoretical framework for the inclusion of students' voices in research. This chapter concludes with examples of research on standardized testing in which students have been included as participants, highlighting studies in which the perspectives of English language learners have been deliberately included.

Federal and State Testing Policies

The current iterations of federal and state testing policy have deep roots in the history of American education: a debate about the importance and validity of standards has been waxing and waning for over a century, with the current wave of standardization and efficiency coming out ahead. In 1912, journalist Joseph Rice called for "a scientific system of pedagogical management [that] would demand fundamentally the measurement

of results in the light of fixed standards" (as cited in Kliebard, 1995, p. 20). Rice argued that educational reform "revolved around a clear articulation of definite goals and on finding the techniques to measurement that would reveal whether those results have been realized" (Kliebard, 1995, p. 20). This ideal is at the heart of modern standards-based reform; reform that has led to increased measurement of student progress towards academic content standards. The term pedagogical management encapsulates the crux of the current dilemma in testing policies – meshing the ideal of ensuring that all students receive a high quality education with the management of that process in a way that is both fair and valid.

Researchers frequently trace the roots of the current accountability movement to 1983 and the publication of A Nation at Risk (Anderson, 2004; Magnuson, 2003; Nichols, Glass & Berliner, 2005; Stevenson & Waltman, 2005). This oft-cited report painted a dismal portrait of the state of education in America, "tying lack of student achievement to impending economic crisis" (Magnuson, 2003, p. 35). Published by the National Commission on Excellence in Education in 1983, the report decried the poor academic performance of American students, and led to "wide-ranging reforms that included an increased reliance on testing, an expansion of the kinds of stakes attached to test scores, and school-level incentives to encourage instruction that would result in higher test scores" (NCELA Assessment and Accountability Resource Guide, p.3). The cry for accountability in education led to a focus on standards, which called for a wider array of assessments measuring more than discrete facts. In fact the National Research

Council[2] describes the standards-based reform model as a theory of action based on four major components: standards, assessments, flexibility, and accountability (2000). Yet a narrowing of the focus of assessment appears underway once again as the fear of sanctions (such as forced school restructuring and offering students choices to attend other schools) coupled with more required testing under NCLB necessitate the need to tap into already limited financial resources in the most efficient manner possible. Though federal law does not prescribe a particular approach to testing, the current political trend is away from performance-oriented assessments that evaluate higher order skills and towards increasingly more uniform measures of student progress in order to meet the requirements of testing more students more frequently in the most cost-efficient manner (Darling-Hammond, Rustique-Forrester & Pecheone, 2005). As a result, reports abound of instruction becoming more narrowly focused on the subjects that are assessed, to the detriment of other curricular areas (Abrams, 2004; Anderson, 2004; CEP, 2006; Stevenson & Waltman, 2005; Sunderman, Tracey, Kim & Orfield, 2004; Taylor, et al., 2003; Wright, 2002). For example, the Center on Education Policy (CEP) reports that of the 299 nationally representative school districts surveyed, 33% reduced the amount of time spent teaching social studies, and 29% reduced the amount of science instruction in

---

[2] Established in 1916, the National Research Council (NRC) is "under the auspices of the National Academy of Science, the National Academy of Engineering, and the Institute of Medicine", part of a non-profit, private institution providing science, technology and health policy advice under a congressional charter signed by President Lincoln in 1863. The mission of the NRC is to "improve government decision making and public policy, increase public education and understanding…" (http://sites.nationalacademies.org/nrc/index.htm).

order to devote more time for reading and math in the elementary grades (2006). In addition to a decrease in exposure to curricular areas not subject to testing, student interest, motivation, and cultural recognition can become lost in the frenzy of accountability testing when instruction becomes closely tied to large-scale assessments driven by the subsequent public reporting of test scores (Gottlieb, 2006).

Rationale

The purported rationale for large-scale assessment policy is multi-faceted. Rivera and Collum (2006), in their introduction to a compilation of recent studies of state assessment policies, identify two primary factors driving the current testing movement in the United States: measurement in order to *identify* gaps in students' academic achievement, and the *inclusion* of all students in that process. The first premise identified by Rivera Collum is founded on the assumption that changes in instruction will occur by measuring achievement gaps and reporting those gaps, which results in increased competition. This is partly fueled by the "flexibility" infused into NCLB – allowing parents the option to leave failing schools. Describing such competition as one of the major principles of the No Child Left Behind Act, Kim and Sunderman (2004) are critical of the theory that competition will result in "better educational opportunities for disadvantaged[3] students and improve the performance of low-performing schools" (p. 5),

---

[3] While the term disadvantaged is used to describe all groups for whom disaggregated data is made available under NCLB, I do not agree with its use in describing English language learners because of the

while Lazarin (2006) claims such competition (fueled by public exposure of test results)

feeds both student motivation and the desire to present a positive report to stakeholders:

> Standards based reform includes three major theoretical components. First, high standards will motivate students to improve their performance if they are challenged by rigorous academic courses. Second, accurate assessments will be used to measure improvement and make important decisions about students. Third, this reform will lead to school system accountability by providing parents, policy-makers, and advocates information about the performance of their local schools. (p. 4)

The assumptions above pertain to all students, including English language learners. In

fact, Rivera and Collum's second factor driving the current testing movement – inclusion

of all students - is noted as a positive consequence of NCLB for English language

learners in several sources (Center on Educational Policy, 2006; Coltrane, 2002; Gottlieb,

2003, 2006; Lazarin, 2006; Taylor, et al., 2003; Wright, 2005). For example, Wright

points out that schools that had previously neglected limited English proficient students

can no longer afford to do so describing this as the "sole positive outcome" of NCLB

(2005, p. ii). In the Center on Educational Policy (2006) survey responses regarding the

impact of NCLB from state (all 50) education officials and interviews with 38 local

district officials, the focus on student subgroups was cited as being a positive outcome.

One respondent indicated how this focus "forces individuals to reflect on their beliefs and

assumptions about the ability of low-income and minority students to achieve." Another

participant said, "We are being held accountable for all students, so we are looking at

---

term's negative connotation. The students are not disadvantaged. Rather, the schools they are in often fail
to make the necessary accommodations to provide them adequate instruction.

everyone! We can't just give them busy work in the back of the room anymore" (p. 7).

NCLB requires the same high standards and expectations for all students (NCLB; CEP,

2006; Gottlieb, 2003, 2006), thereby creating a shared sense of responsibility among all

educators for the education of all students.

Lachat, writing in 1999 (pre-NCLB), described a direct link from accountability

to better teaching and referred to the accountability movement has having "the leverage

needed to raise expectations for English language learners, and the emphasis on higher

level skills should improve the quality of teaching provided to them" (p. 60). Rivera and

Vincent (1997) refer to the same rhetoric of inclusive measurement. They assert that by

including ELL in assessment systems, they can also benefit from any reforms that result

from assessment data. Acknowledgements of the benefits of the recent, inclusive

requirements for English language learners are a tacit acceptance of the fact that prior to

such inclusive accountability requirements, the academic needs of English language

learners were not adequately being met.

Others have cautioned against mere rhetoric of inclusive measurement. Lacelle-

Peterson and Rivera (1994), when writing at the cusp of assessment reform and the shift

in emphasis from testing to a broader definition of assessment, prophetically opined that

if reformers do not seriously consider the implications for specific groups of students,

among them ELL, little change will occur:

> Though the motivating impulses and underlying commitments of the current
> standards-setting and assessment reform efforts are profoundly optimistic and
> egalitarian, the processes and results have neither considered nor adequately
> addressed the perspectives and needs of ELLs. The implicit guiding assumption

appears to be that whatever curricular revisions and/or assessment innovations contribute to the success of monolingual students will also work for ELLS – that once ELLS know a little English, the new and improved assessments will fit them too. (p. 56)

Gottlieb (2003) concedes the benefits of demanding the same level of rigor for ELL, as long as appropriate instructional modifications are made. She optimistically describes how teachers now have "ammunition to insist that the same rigorous, content-based curriculum be used with all students while using instructional strategies that take into account the unique characteristics of English language learners" (p. 8). Her assumption is that instructional strategies are, in fact, being used. While teachers may now have "ammunition", there is still the need for reasonable accommodations not only in teaching strategies, but in approaches to assessment as well.

A primary rationale for the inclusion of all students in accountability systems, including ELL, stems from the prior exclusion of some students and the resulting lack of need for schools to be accountable for their learning, or lack thereof (Anderson, Liu, Swierzbin, Thurlow, & Bielinski, 2000; Lacelle-Peterson & Rivera, 1994; Lazarin, 2006). Yet current testing policy fails to acknowledge the challenges Lacelle-Peterson and Rivera identified more than a decade ago – the necessity to consider the unique instructional and assessment needs of all students. Regardless of how and whether teachers use assessment data to make necessary changes in instruction, the complications lie with the measurement process itself: Does the process that works for monolingual English speakers work for all students? With NCLB, accountability has resulted in

sanctions, and resulting reforms are often reactionary and fear-based, rather than focused on sound decisions made on the basis of local knowledge and student characteristics.

Local knowledge and input has become decreasingly emphasized in the shift towards federal control over education in the U.S. This shift has been identified as a foundational underpinning of current K-12 assessment policy. In a global sense, Shohamy writes that the primary role of high-stakes tests has shifted "from measuring knowledge to enabling centralized bodies to control education, content, learning, and teaching" (2001, p. 36). She also describes how tests can be a "disciplinary tool"; perceived as being authoritative and having the power to dictate educational policy that affects society as a whole. With the locus of control at the test, knowledge itself is redefined by the parameters of tests. The narrowing of curriculum to that being tested, and the focus on discrete facts in standardized tests are examples of this phenomenon. Citing the "unprecedented federal intervention on state-level education policy making that directs all states toward a single goal (i.e., 100 percent of students reaching 'proficiency') via a single system of implementation (i.e., standards-based assessment and accountability)", Nichols and her colleagues elaborate on the control involved with our current testing system, claiming that its historical base can be found in the rhetoric of a failed education system internalized by Americans since A Nation at Risk (Nichols, Glass & Berliner, 2005, p.4).

With a new rhetoric of standardization and control comes a new accountability for policy makers themselves. Spillane, et al. (2002), writing at the dawn of NCLB, note

25

"the legitimacy of policymakers flows from the consent of the governed" (p. 388), and the "governed" typically prefer some discretion in implementation – especially local administrators and educators. They note "[d]uring the past 50 years of educational policymaking, this dialectic of requiring change and allowing for local autonomy has played itself out" (p. 388). When compliance is the goal of policy, tensions rise. Magnuson encapsulates the tensions between policymakers and local educators; it is politically important to test every student to "allow policy makers to aver that they are indeed concerned about every student" (2003, p. 23). He contrasts the "tremendous breadth of accountability" under NCLB with the lack of "depth of accountability"; results have little direct use for individual students or instruction given the delay in obtaining results. Yet policymakers are looking to leave their positive imprint on educational reform, which requires "measures that allow easy comparison, that generate improvement quickly, and that are politically viable, meaning, in part, not prohibitively expensive" (p. 23). Public perception plays a key role, primarily because the tenets of NCLB – to set benchmarks for achievement, test progress, and punish schools where students fail - are "plausible and easily understood by the public" (Crawford, 2004, p. 2). There is a pervasive belief that exposure (data reporting) will result in public pressure to reform, and sanctions tied to student performance have upped the ante. Prior to NCLB, assessment policies were "more carrot than stick" (Crawford, 2006), and states had much more discretion in determining the nature and volume of large-scale assessments.

Given that discretion for current testing policy implementation at the state level is largely superficial at best, it behooves us to consider the agents at the local level. The authors in the aforementioned reviews of testing policies attend to the rationale for such policies from a societal point of view. In all cases, students are assumed to be the passive recipients of well-intentioned, though flawed policy, yet rarely merit recognition as stakeholders worthy of consideration, let alone agents in the implementation process. The good intentions of those instituting a policy are often sidetracked when the policy intersects with the actors in a school (Abernathy, 2007). Explorations of policy implementation have traditionally focused on the agency of principals and teachers in the implementation process (Abernathy, 2007; Seashore Louis, et al., 2005; Spillane, et al., 2002). The opportunity to develop shared understandings and explore multiple interpretations of schooling and educational policy from students' perspectives is a focus of a small but growing body of research. A rationale for such research and a theoretical framework which builds on that used to explore teacher agency in policy implementation are described in the next section.

Theoretical Frameworks

Multiple theoretical frameworks serve as a rationale for conducting research that engages students in meaningful and reciprocal learning opportunities regarding educational policy: Sense-making, Constructivism, and Critical Theory. The same multiple theoretical constructs (or a combination thereof) inform much of the research

27

involving students' perspective and voice research (Cook-Sather, 2002; Cook-Sather, 2006; Lincoln, 1995; Oldfather, 1995). In advocating for a multiple theoretical approach to research, especially with children, Graue and Walsh (1998) note how a dominant theory "…can become hegemonic, dictating not only how to see the world, but also what to look at and what not to look at" (p. 26). Rather than a single theory which functions "like a set of blinders", multiple theories allow for multiple road maps and the opportunity to reflect on phenomena from various angles. The multiple road maps that inform this study are described in the following paragraphs.

Constructivism is a theory which views knowledge as constructed by students, not merely discovered. It advocates student autonomy and initiative in constructing their own knowledge, with adults sharing responsibility in that process (Cook-Sather, 2006; Pinar, Reynolds, Slattery and Taubman, 2004; Balderrama & Diaz-Rico, 2006). Lincoln (1995) advocates for a constructivist approach to research when working with students because of the recognition of "multiple realities" inherent in this approach, leading to the construction of those realities grounded in the "particulars of personal experience" and "the webs and patterns of influence that operate on individual lives" (p. 92). Jill Koyama (2004) offers a description of this in practice, noting how federal and state testing policy becomes appropriated, thereby delineating "positions for ELL students to inhabit", and how "students and teachers, together, negotiate and give meaning to these positions by selectively implementing chosen elements of the policy and excluding others" (p. 403). Such positions are complicated for all students, but particularly for an English language

28

learner, who is "never independent of the ways in which other social actors construct and express an interpretation of him or her..." (p. 404). Teasing out these positions becomes a process of co-construction which relies on a culturally sensitive, as well as a constructivist approach (Tillman, 2002). Philosophically, constructivism resonates with the stance of this research in that students do not attend school or take tests in a void. Therefore, the meaning they bring to an event such as large-scale testing is negotiated on several levels: among their peers, in dialogue with teachers and parents, and through the very act of participating in the research described in this study.

The meaning-making activities of constructivist research parallel the sociological theories of sense-making, which have primarily been applied to teachers. Schmidt and Datnow (2005) refer to symbolic interactionists such as Blumer (1969) to describe how "human beings act toward things on the basis of the meanings they have constructed for them". They say that "…meanings are modified through an interpretive process used by the people involved in those encounters. Sometimes meanings are contested, or they are affected by power relationships in a given interaction" (p. 950). Prior knowledge, beliefs, attitudes and experiences are important components in the differentiated interpretations of policy. Spillane, et al. (2002) have created a 3-staged framework for understanding sense making. The tiers include a) individual cognition (including prior knowledge, beliefs, experiences and emotion), b) situated cognition (where multiple dimensions of a context or situation play a key role), and c) the role of representations or policy stimuli. In short, policies do not maintain uniformity on the road to implementation. In fact, local actors

have input into the ways in which a policy ultimately gets interpreted, as described by

Spillane, et al. (2002):

> Policy messages are not inert, static ideas that are transmitted unaltered
> into local actors' minds to be accepted, rejected, or modified to fit local
> needs and conditions. Rather, the agents must first notice, then frame,
> interpret, and construct meaning for policy messages. (p. 392)

The cognitive, or sense making dimension of the implementation process provides a

logical frame for a further exploration of student voice in educational policy. Though

most work in this area focuses on teachers as the implementing agents, (Abernathy, 2007;

Spillane, et al, 2002; Seashore Louis, et al., 2005), the possibility exists to extend such

explorations to students and the sense they make of policies that have a direct and long-

lasting impact on their futures. Rather than situating the locus of control for a policy's

implementation at the teacher level, this research extends this work one step further,

recognizing students as the ultimate actors in the implementation process.

The work of Alison Cook-Sather (2002; 2006) adds a critical theoretical

perspective to the constructivist and sense-making ideas that students need to be both the

"authors and assessors" of their learning, noting how the foregrounding of the political

nature of education can lead to negotiated learning about social injustices and inequities,

ultimately empowering students. Often times, students can clearly articulate the injustices

and inequities that surround them, but are seldom given the forum to do so (Nieto, 1994),

or the chance to explore the root causes or systemic complicity underlying those

inequities. In general, large-scale testing creates a myriad of implementation scenarios

ripe for inequities, where decisions are made that directly impact students. In the case of

this research, that impact is compounded by the fact the participants are English language learners, primarily "of color", and representative of a variety of linguistic, religious, political, and cultural backgrounds often ignored in structural contexts of public education.

Lincoln (1995) notes how elements of critical theory have great relevance to research with students, namely a focus on the structural elements of research participants' contexts. Further elements of critical theory that bolster the search for students' voices include the opportunity to help students examine patterns in their lives and to discover the somewhat hidden structures that shape them. Lincoln claims that students do not need "doctorates in cultural ideology to uncover the meta-messages of the larger society" (p. 92), but rather a chance and a forum in which to do so.

Simply creating the forum in no way guarantees the emergence of student voice. Describing the "arduous and complex process of coming to voice," Donaldo Macedo (2006) cautions against "facile propositions" about giving voice to others:

> …voice is not something to be given by those in power. Voice requires struggle and the understanding of both its possibilities and limitations. The most educators can do is to create structures that would enable submerged voices to emerge. It is not a gift. Voice is a human right. It is a democratic right. (p. 4)

Fielding (2001) adds a caution against simply adopting the "rhetoric of student voice", suggesting the need for "new communities of practice shaped by an essentially dialogic form of engagement" (p.108) in order to engage in reciprocal learning with students. Cook-Sather (2002) offers an additional caution from a poststructural/feminist standpoint

– the need to be careful not to "uncritically or unreflectively [privilege] student voices" (p. 6), but rather, to consider the many intersecting variables (language, context, power, etc.) and the dimensions of sense making that influence voice.

Seeking students' voices, then, becomes a deliberate act of recognizing their agency and a validation of the sense making process. To do so requires a cultural shift, a "repositioning of students in educational research and reform" (Cook-Sather, 2006, p. 365). As mentioned earlier, policy implementation research has focused on the teacher as implementing agent. This study is built on the premise that students have the final say – perhaps not as 'implementing agents' so much as 'enacting' agents. As intended recipients of testing policies, they have the ultimate say in the enactment of those policies.

The following section explores student voice research in more detail, beginning with a historical perspective and an exploration of theoretical orientations within voice research.

Student Voice

The study of students' experiences in schools has grown considerably in the last decade. Terms used to describe research in which students play a central role range from student experience to student voice research. There are many commonalities among the work of researchers exploring student perspective, experiences, voices, etc., yet varying

ways in which they describe the progression of student involvement. Key contributors to this type of research and their respective typologies will be discussed here.

In his introduction to the International Handbook of Student Experience in Elementary and Secondary School, Dennis Thiessen describes student experience research as an evolving field, paralleling the growth of childhood studies (an interdisciplinary field bridging the disciplines of sociology, history and policy study related to children's rights) (Thiessen & Cook-Sather (eds), 2007). Thiessen contrasts studies of student experience with studies of children, noting how studies of student experience "critically examine the individual and collective worlds of students, with a particular focus on how they navigate and negotiate the dynamic and multi-dimensional demands of their classroom and school lives" (p. 5). Student experience research encompasses a variety of distinct, yet interrelated orientations (Thiessen, 2007), including:

1. How students participate in and make sense of life in classrooms and schools;

2. Who students are and how they develop in classrooms and schools; and

3. How students are actively involved in shaping their own learning opportunities and in the improvement of what happens in classrooms and schools (pp. 5-6)

Thiessen traces the historical roots for such research to the ongoing, longstanding debate representing dichotomous views of students in school:

one within which students are seen as unknowing neophytes who benefit from the wisdom passed on to them by knowledgeable teachers (transmission image), and the other, favoured by educators and scholars who advocate the importance of school experience, where students are capable and active agents in their own development who benefit from a nurturing and enabling learning environment (discovery-based image). (p. 6)

He goes on to use the double entendre of "knowing students" (p. 53) to encapsulate student voice research. Not only are researchers, teachers and other adults interested in learning from students – there is also recognition of students' sense making processes as well as their agency in matters of schooling. This is reflected in the shift away from simply studying students' attitudes and perceptions towards studying students' perspectives and voices. Elaborating on the distinction, Thiessen describes perspectives as "constructed, individually varied, situated or contextually bounded, and negotiated in the socio-political realities of classroom and school life" (p. 54). Noting that student perspective research illuminates nuanced, "more complicated and changing portrayals of how students make sense of their lives at school", and in conducting studies of this type, "researchers recognize that students have both authorship of and authority in their lives at school" (p. 54). They have valid perspectives, and they have voice.

Cook-Sather (2006) explains the emergence of the term "student voice" to describe a range of activities that aim to "reposition students in educational research and reform" (p. 359). Citing Holdsworth (2000, p. 355), she explains how the term "voice" signals having a legitimate perspective and opinion, being present and taking part, and/or having an active role in decisions about and implementation of educational policies and practice" (Cook-Sather, 2006, p. 362). Depending on the context, voice can range from

34

the sound of students speaking, to having real presence, power and agency regarding issues that affect them (Cook-Sather, 2006; Holdsworth, 2000).

Rudduck (2007), further delineates the difference between perspective and voice, noting that those interested in perspective (referring specifically to teachers) "may see students primarily as sources of interesting and usable data, but they are less likely to have goals that are expressed in terms of community" (p. 591). Mitra (2007) refers to this as the first of three types of student voice research – that of adults *listening* to students. While listening to student perspectives can provide great benefits and inform change, students have little ownership in this type of research, nor does it guarantee change in the status of students (Rudduck, 2007; Mitra, 2007).

For those concerned with 'voice', students' status – especially if marginalized -- becomes a factor. Research and initiatives in this vein, "usually take on the serious and significant task of eliciting and presenting the experiences and views of groups on the margins, thereby helping them to move from silence and invisibility to influence and visibility (Rudduck, 2007, p. 591). Doing so involves a more collaborative relationship with students, one where ownership is shared (Mitra, 2007), and students are treated respectfully and taken seriously, as indicated by the following quote:

> If [students]suggestions are acted on, or if there is some explanation of why they cannot be acted on, then [they] are more likely to feel that they have a stake in school and they are more likely to commit themselves to its learning purposes. Second, being consulted and knowing that what you say is taken seriously builds students' self-respect and gives them a sense that others respect them and this, in turn, can also strengthen their commitment to learning. (Rudduck, 2007, p. 599)

Dana Mitra (2004) found that including students in a major reform effort led to increased agency, belonging, and competency – assets central to youth development. Her case study explored two student groups involved in a three-year high school reform effort. One group was comprised of Latino students (half of the school's population was ELL) who worked at student level initiatives, namely tutoring and translation assistance to other students. The other group, Student Forum, focused on organizational level activities – "building communication and partnership schoolwide between students and teachers" (p. 657). As an outside observer, Mitra attended meetings and interacted informally with staff and students, and interviewed students, teachers and administrators. Her data collection occurred for two years, allowing her an in-depth look at the student voice activities that were happening at this school. And while she was not a participant in the activities, she was a frequent observer at the school, and offered her notes back to the groups for member checking, as well as to give them a record for their own purposes. Mitra found "little discrepant evidence" (2004, p. 660) contradicting the strong increase in student agency, belonging and competence among student participants in each of the groups.

Fielding (2001) alludes to the distinction of students as sources of data or as agents of transformation. Are they stakeholders worthy of consultation (an important first step and the least threatening) (Mitra, 2007), or are they brought in as important decision-makers regarding *their* education? Ultimately and ideally, student agency evolves beyond shared decision making and toward leadership. "If increasing student voice truly means

36

sharing the ownership of school decision making with students, then youth must do more

than speak their minds about problems; they must have the opportunity to lead the way

toward innovative solutions" (Mitra, 2007, p. 743). Figure 2.1 provides a graphic

representation of the various typologies of student participation in educational reform.

Figure 2.1: Typologies of "voice" research.

Holdsworth (2000, p. 358). Student Participation Ladder.	Thiessen (2007)	Mitra (2007)	Fielding(2004)	Current Study
'Speaking out'				
Being heard	Attitudes / perceptions			
Being listened to	Listening	Students as data sources	Sense-making
Being listened to seriously and with respect	Perspectives			
Incorporating youth/student views into action taken by others		Collaboration	Students as active respondents	
Sharing decision-making, implementation of action and reflection	Voices		Students as co-researchers	
		Leadership	Students as researchers	

Regardless of how these key contributors to the field of voice research position

themselves, at the core they have common commitments:

> …to listening and responding to a diverse set of perspectives and not just
> tolerating or tokenizing them but always destabilizing the center; to
> acknowledging that what you don't know is much bigger than what you
> know; to the notion that the project of school is an ongoing negotiation
> rather than transmission; to the idea that education is a process based on
> rights and relationship; to the most basic premise that education is about
> change. (Cook-Sather, 2006, p. 382)

To that end, the incorporation of a sense-making framework into the current study moves slightly beyond listening, or simply viewing students as sources of data. Rather, a deliberate attempt was made to engage in dialogic reflection, in collaboration with students, and to not simply hear them, but to listen to their perspectives.

Overall, the need for youth to have opportunities to lead and be involved in school change is becoming increasingly apparent, as educational systems remain largely static, while each subsequent generation of youth is changing rapidly. Rudduck (2007) notes how, "Over the last 20+ years, schools have changed less in their regimes and patterns of relationship than young people have changed" (p. 588). In addition to providing a fresh perspective of reforms in action, student voice research alters the very status of students in schools. Historically, support for such research has come from various perspectives or paradigms, including human rights (specifically stemming from the 1989 U.N. Convention on the Rights of the Child), citizenship, or from those "who endorse the perception of young people as consumers and who argue that as a key group of stakeholders in schools they should have the chance to comment on their schooling" (Rudduck, 2007, p. 589). Each of these perspectives is embodied by critical theorists, as well.

In their examination of school life for refugee and immigrant students, Amy Roberts and Steven Locke (2001), paraphrase critical theorists; "Student voice can become an organizing force to negotiate and construct multiple interpretations of school life within the reality of institutionalized ways of being in school" (p. 375). As Nieto

(1994) reminds us, "students spend more time in schools than anyone else except teachers [yet are] often given the least opportunity to talk" (p. 420). But voice is not something that can be given by someone, and research in this vein requires attention to the nuanced complexities that comprise the lives of adolescents in school.

Special Considerations in Researching Student Experience

Listening to participants is predicated on a number of crucial variables, which also impact the methods used to elicit data:

> Students may not know they have a voice. Or if they know they have one, they consciously repress it when in the presence of adults in authority. Consequently, patience in the process of letting student voices emerge is essential, as is active listening, probing, and…nonjudgemental brainstorming. (Lincoln, 1995, p. 90)

Patience is important, not only for allowing a trusting relationship to develop, but also for the passage of time, especially when researching with students who are relatively new to the country or to a particular setting. Immigrant students have complex views about education and the nature of schooling that sometimes differ from those espoused by U.S. schools, and Valenzuela (1999) describes how many have a "dual frame of reference" ; comparing their present status and attainments to "typically less favorable situations 'back home'" is a form of silencing. Referring to students in her study, she notes how "their sense of being privileged to attend secondary school saps any desire they might have to insert their definitions of education into the schooling process" (p. 14). Sense-making for students such as those in Valenzuela's study requires time to reflect on previous experiences with school and to develop a level of comfort with critically

examining their current educational opportunities – in other words – develop an

awareness of their own power and agency in their new context.

Schultz (1997) describes the real and imagined boundaries she needed to cross as

a white, middle-class researcher in order to gain the trust of her participants (female

students of color). Her two year study included observations, writing samples and

interviews with urban students to get their perspectives on their literacy practices in and

out of school. While her primary role was as researcher in the project, she positions

herself as an educator in her article, as well as at the inception of the data collection.

Having a respected teacher introduce her to the students as a friend put her immediately

into a category of one to be respected by affiliation, though many other boundaries had to

be crossed in her research, spurring her to reflect on the "blinders" researchers bring to

the process, and the power we exercise without even realizing it. Schultz gives the

example of realizing just how closely students in her study had been observing and

examining her, while she was doing the same to them. They had noted everything from

her affiliation with their respected teacher, to the type of car she drove.   Constantly

working on creating a reciprocal relationship with her participants and maintaining a

close level of involvement, she notes:

> …intimacy implies responsibility that extends beyond the research project.
> Although I was aiming for a reciprocal, participatory, and respectful
> research project, I was always aware that the questions and authority were
> my own. It was their story I was collecting and writing. (p. 500)

 Schultz acknowledges the marginal status she maintained in her participants' lives, and

the agency they maintained throughout the process – they ultimately chose which stories

to share. As researchers, "…we operate on the margins. When we cross boundaries in teaching and research, we need to build a different kind of trust at the same time that we remain aware of the responsibilities this trust implies" (p. 502).

Others have built on Schultz's work by examining adult agendas and the adult interpretive frameworks that often exacerbate the adult/adolescent dichotomy (Cook-Sather, 2002; Silva & Rubin, 2003). Cook-Sather suggests a need to carefully consider our tendency to make sense of adolescents' worlds within our own analytical framework, and highlights work by Oldfather and others, where students have been invited to be co-researchers by helping in the analysis, writing and presentation of findings. This, however, requires a paradigm shift, as it is a "psychological challenge" to convince policymakers, teachers, researchers, and administrators that they have something to learn from listening to students. In describing the need for adults to reconceptualize our assumptions about "who can and should be an authority on educational practice", and move beyond the rhetoric of student voice Cook-Sather adds:

> It is easy to assert that a first step toward including student perspectives on schooling is counting students among those who belong on the list of stakeholders with a voice in shaping educational policy and practice. It is much harder to actually change the ways we as educators and educational researchers think about and interact with students. (p. 9)

Increasingly, researchers are taking up the call for allowing for the emergence of student voice, either nominally by allowing youth to "share their opinions of problems and potential solutions", thus validating their perspectives. Further along the continuum,

41

voice research involves "young people collaborating with adults to actually address the problems in their schools" (Mitra, 2004, p. 651).

High school-aged students, in particular, have the cognitive and emotional maturity coupled with the desire to analyze their experiences critically (Roberts & Locke, 2001). In an article exploring three refugee and immigrant students' experiences (as a subset of a larger, comprehensive qualitative study), Amy Roberts and Steven Locke draw on critical theorists in an attempt to find and develop student voice. Through the telling of their life history stories, students were able to interpret their school experiences and describe the strategies (both cognitive and affective) they employed in their daily interactions in school. Doing so helped shape their personal identities as students, and underscore the importance of viewing each student's life story as unique and influential in their overall academic success. The researchers emphasize the responsibility of educators to "accommodate to the students, not the other way around" (p. 390), as well as the need for students to take responsibility to pursue an active and interactive education, putting themselves into the system of school life.

As preparation for becoming productive citizens in a democratic society, others have emphasized the importance of giving students the opportunity to participate in *decision-making* in high school about aspects of school that affect them daily (DiMartino, as cited in Miles, 2004) or to embrace a critical pedagogy that has a "political goal of creating a more just society instead of accommodating and adjusting students to the existing social setting" (Moraes, 1996, p. 107). This requires students to be actively

42

engaged rather than passive recipients of knowledge - to prepare students to leave school "thoughtfully engaged and challenged rather than indoctrinated" (Roberts & Locke, 2001, p. 391). But as Macedo (2006) writes,

> There is no time for such reflection on the part of teachers and students – they are much too busy running a maze of technical procedures. The industrial model of school requires students who are good team players, respect authority, value the work ethic, follow directions, and are passive. (p. xv)

Referring to schools as places where information is sent "down the industrial conveyor belt into students' minds" (p. xi), Macedo criticizes the lack of opportunity for students to engage in transformative pedagogy. Corbett and Wilson (1995) describe how this "ingrained passivity" among students is at odds with constructivist approaches to learning (p. 13), noting contradictions between what students expect and the expectations of teachers, administrators and others. Recognizing the value of research inclusive of student expectations and perspectives to inform policy makers and practitioners, Yonezawa and Jones (2007) examine the particular relevance of such research to testing. "Student voice can augment standardized test scores, which for many districts and schools have become the ultimate marker of educational success and failure" (Yonezawa & Jones, 2007, p. 681). The voices of students can, in fact, provide a more nuanced picture of the validity in ascribing such weight to test scores.

Examples of research in which students, nominally, were counted as stakeholders with regards to standardized testing are explored in the next section. It begins with a global exploration of research involving high school students and testing, followed by those studies with either emphasis on or deliberate inclusion of English language

43

learners. Not only have students' opinions traditionally been peripheral to many aspects of schooling, the voices of marginalized populations within the student demographic have been even more so. This research is an attempt to learn from one particular subset of the population that is often marginalized – English language learners, whose background experiences as well as cultural and linguistic diversity create unique considerations for research under the theoretical frameworks identified above.

Student Perspectives of Standardized Testing

Using Sutton and Levinson's (2001) conceptualization of policy as a "complex social practice, an on-going process of normative cultural production constituted by diverse actors across diverse social and institutional contexts" (p. 1), the inherent complexities of federal and state educational policies become glaringly obvious if examined from the point of view of the students. The ways in which policy has been appropriated (Koyama, 2004; Sutton & Levinson, 2001) varies according to how "local actors [interpret and adapt] such policy to the situated logic in their contexts of everyday practice" (Sutton & Levinson, 2001, p. 17). In other words, how do actors (students, in this case) make sense of testing policy? While the authors do not refer specifically to students in the introduction to their edited text, they emphasize their view of policy as more than simply a mandate – it is a "contested cultural resource" impacted by the "purposeful practice of diverse social actors [reinstating] agency across all levels of the policy process" (p. 3). One cannot overlook or underestimate the key role that students

play in the process, especially when testing policies embody high stakes. And the stakes are most likely to be higher for *students* (as opposed to high stakes for schools or systems) in high schools.

Research studies in which high school students are asked their perspectives of standardized testing programs are scarce. Where studies do exist, researchers often approach students as mere sources of data, failing to evoke even the rhetoric of student voice. Many utilize survey data or are limited to those ethnographic studies that contain relatively few examples of actual quotes from students, where "voice" may have been intended but failed to emerge because of how the student data were represented. Existing studies can also be categorized as one-shot reflections - brief snapshots in time - versus those that engage students more critically and long-term.

The studies that follow are examples of researchers listening to students – an important first step on a journey towards allowing students an opportunity to engage with and impact policy.

Perspective Studies

The Texas educational accountability system has a well-established tradition of testing and has been the focus of a variety of studies where students have been asked to reflect on their testing experiences in both the K-12 and post-secondary systems. A comprehensive collection of students' opinions about a statewide, high stakes exam was done at Texas A & M University Corpus Christi in a freshman writing class (Blalock &

Haswell, 2003). Responses to the following optional writing prompt were gathered in the fall of 2001 and 2002:

> *What was the TAAS [Texas Assessment of Academic Skills) experience like for you as a student in English classes? Describe your activities and feelings as you prepared for these exams, took them, and learned of the results. Overall, was it a good educational experience for you? If you didn't do TAAS, you can write about any standardized test preparation and testing that you have experienced in school--SAT, AP, etc.*

The purpose of posting the responses on a web site was to let the opinions of students be shared, though the authors of the information note that only students who passed the TAAS were included in the sample (a requirement for enrollment in a university course). Of the 854 responses over two years, 63% were completely negative, and only 8% were completely positive, with the rest of the responses falling somewhere in between. A common thread throughout the responses was the inordinate amount of time spent preparing for the TAAS throughout high school, with one student referring to the preparation as "boot camp". Several others noted that for strong students the test prep was too easy, shortchanging them of advanced curriculum. For these college students, their overwhelmingly negative reactions have to do with the fact they were precluded from receiving college preparatory educational experiences in high school because of the tunnel-like focus on the TAAS. The voices of students who had not passed this test are absent from this data, leaving one to speculate on the extent to which they, too, felt short-changed by their high school educational opportunities.

While the previous study invoked retroactive reflections, another study by Flores and Clark (2003) involving the Texas accountability system asked students in grades 3-11

to respond in writing to the following question: *What do you think/feel about TAAS?*

These students, representing a variety of geographic and socioeconomic backgrounds, had perceptions that appear to vary based on grade level (at the time of the study, younger students were not yet faced with grade retention based on test results). Recognizing that teachers and students "are the ones engaged in the daily realities of testing" (p. 3), the authors consider them important stakeholders in the debate about standards and accountability, even adding pre-service teachers to the category of stakeholders worthy of having a "forum [in which] to speak out". Student responses indicate that beginning with 5th grade, there is an increasing concern about not passing the test and the potential impact on future education and occupation. Middle school students were less concerned with their performance, but felt the test was: "Dumb, because we didn't know why we have to take it", "stupid", "has no point", "boring", and a "waste of time". One particularly telling comment that would be worth a follow up study was: "After taking the test, we shouldn't have any more school; there is no point to it" (p. 19). Their responses speak to the lack of investment they have in the testing process, or even school itself if they saw no point to going after the tests were over. Interestingly, by high school many students in the study felt the standards of the tests were too low. Their reactions ranged from annoyance to being concerned enough to spur some degree of motivation. Overall, the results of this study show that students do not have a clear understanding of the purpose of testing or what it means to pass the test. The researchers note "the existence of so many differing perceptions and concerns about the

test calls into question the validity of TAAS results. Yet, schools and teachers are judged by these results" (p. 21). In both of the examples described above, important student voices were shared – yet the methods of gathering responses leave questions unanswered. In the K-12 study, researchers had no direct contact with students (Flores & Clark, 2003), and neither example allowed for any extended conversation or follow up discussion about the responses. It is also unclear in both examples as to whether any of the participants were English language learners, and what impact, if any, this had on their experiences. Regardless of any perceived shortcomings in studies such as those above, the reported misunderstanding of and lack of engagement with the testing process are points not to be taken lightly, and a theme that emerged with students in other states.

A study by Wheelock, Bebel, and Haney (2000) reveals a wide range of student responses to testing, yet took a different twist in terms of research methods; students were asked to draw self-portraits after taking the Massachusetts Comprehensive Assessment System. Their portraits showed a range of affective responses: some as diligent, confident thinkers and problem solvers but a greater number as anxious, angry, bored, pessimistic, or withdrawn. The authors concluded that students' beliefs, attitudes, and feelings about testing may influence testing behavior and scores. Noting that students' responses "are rarely factored into decision making regarding high stakes testing policies" (p. 1), the investigators are clear in their assertion of why student responses *should* be included as a factor. "If nothing else, the range of students' responses to MCAS reminds us that high stakes testing policies may have little meaning if, in the process of taking such tests, a

sizable group of students become disenchanted with learning" (p. 6). They concluded that high stakes alone don't necessarily motivate students, especially as students get older. It is unclear, however, if students were asked to reflect on their portraits, or if they were interpreted by the researchers, in which case adult lenses may have uncritically assigned value to student perspectives. Even though the implications are crucial from studies such as this one that question student engagement with the testing process, they remain largely unexplored as do the implications of policies that generate disenchantment with the process of learning.

If students are not motivated during the testing process, how *cooperative* can they be expected to be regarding the intent and implication of the overarching policy? Melissa Chabran (2003) explored the direct role students play in the process of testing in four states; two where the stakes are high and direct for students (tied to graduation), and two where scores are used only for system accountability (but tied to school pride and community, therefore indirectly high stakes for students). Describing how "ultimately, the effects of accountability policies depend on the cooperation of the students", she questions whether "many state assessments are capturing what states expect students to know with reliability and validity" as well as whether an assessment is inherently motivating for students (p. 129). In focus groups with students (not ELL), Chabran found high-stakes tests are anxiety producing and intimidating. One student described test results as a matter of luck, noting how he could either pass by guessing or fail if the material tested was something he had never been taught. Another student preferred exams

to portfolios even though she had not yet taken the Regents Exam at her school, while yet another New York student would prefer a Regents exam "that required them to show their work [as opposed to multiple choice]". In addition to having a concern over individual performance, students in the focus groups believed that school reputation (and therefore overall high test scores) was important to colleges, especially if they were: a.) minority, b.) from NY, c.) went to public school, and d.) had low SES (p. 141). Teachers in the other states talked about students not taking accountability tests seriously, especially if there were no high stakes (tied to graduation) for students.

Overall, the author found that when stakes are located elsewhere (i.e. school accountability), it is difficult to motivate students, especially those students who are not strongly connected to the school in the first place and, as one teacher noted, the delay in reporting of scores is so great that students have often moved to the next grade level before scores are available, essentially rendering them meaningless to students and teachers alike. Chabran acknowledges that her overall sample of student responses is small, but offers that her study "provides insight into what researchers and policy makers might ask students", because responses indicate that "students are thinking quite meaningfully about how accountability policies in their respective states affect them" (p. 144). Chabran reminds us of the need to examine the ways in which students appropriate policy and the degree to which that process is motivating for them. Her study and others described here are important contributions towards documenting a pattern of students' beliefs, experiences, and overall perspectives of tests for which stakes can be high.

Whether high stakes tests are motivating to students or not may have a lot to do with background characteristics as well as the size of the learning gap to be overcome. Roderick and Engel (2001) describe results of a study that somewhat contradicts Wheelock et al. (2002) and Chabran (2003) in terms of student motivation towards high stakes testing. Roderick and Engel studied the effects of a policy to end social promotion in the Chicago Public school system. They interviewed 102 low-achieving 6th and 8th graders (African American and Latino) to ascertain their opinions of a policy to end social promotion based on scores on the Iowa Test of Basic Skills. Did the Chicago public schools' policy lead low-achieving students to work harder? Did it impact their learning goals? The majority of students responded positively; they were creating incentives through goals, were provided a lot of feedback and noted the importance of the classroom context and a caring teacher in helping them achieve their goals. "Helping adolescents understand the policy, making them feel supported and efficacious in achieving their goals, and structuring their activities so they are meaningful all seem to be essential components for success" (p. 221). In all of the research reviewed, this was the only explicit reference to helping students understand policy and the importance of aligning students' goals with the policy, if in fact its intentions are going to be met. The researchers stress, however, the importance of building teacher capacity to meet the needs of low achieving students, noting that testing policies themselves are only "one piece of the puzzle". They also offer several other cautionary remarks, noting that test score gains could represent a number of phenomena: actual gains in learning, effective test

51

preparation, or focused attention during act of testing. They also note that the promotional criteria under the policy were "quite low", with cutoff scores for being promoted to high school being one and one half years below grade level, which might "give students a false sense of accomplishment" (p. 221).   Though it appears that the threat of retention was motivating for some students, the authors caution that "policies that rely heavily on student motivation to improve achievement may place students with the lowest skills in a very difficult position. These students face the greatest task and at the same time often have the fewest resources to accomplish that task" (p. 219). Teachers of the focal students often misjudged them as not caring, rather than recognizing that their low-motivation may be tied to their low skills and lack of external support. Those with potentially the largest gaps -- newcomer English language learners (those in bilingual programs 3 years or fewer) -- were excluded from the study, leaving a description of their intersection with this policy lacking. The researchers conclude their discussion with a final cautionary interpretation of their findings – students' lives are complex, and while policies often raise issues of inequities, they just as often fail to address the root causes of those inequities.

The preceding studies demonstrate how student input can expose a variety of additional factors that affect the way in which testing policy becomes appropriated. In some cases, students do not have a clear understanding of the purposes for the tests; in others, motivation impacts their degree of cooperation during the testing process, which

52

in turns raises questions about the validity of the scores and the stakes to which those

scores are attached.

Perspective studies in Minnesota

In cases where the purpose of the testing was relatively clear, students were able

to give informed decisions about possible improvements to the testing process, and

appreciated the chance to reflect on the process with researchers. Data from two

Minnesota studies that sought students' opinions of the Minnesota Basic Skills Tests

(BSTs)[4] of Reading, Math and Writing reveal that students had a clear understanding of

the purpose for the tests and had keen insights to share about the process (Quest, Liu &

Thurlow, 1997; Thompson, Thurlow & Walz, 2000). In the first study, 96 high school

students with disabilities (not ELL) were asked by Thompson, Thurlow and Walz (2000)

to describe how many times they had taken the BSTs, what accommodations they used,

whether those reflected the accommodations used in their classes, and what other

accommodations, if any, they felt might be useful in the future. The authors reflect that

"common sense is beginning to prevail as people realize that, disability or not,

adolescents seldom follow directions without question" (p. 5). The authors were able to

document that nearly all students interviewed had knowledge of the BSTs, knew how

---

[4]Passing the Minnesota Basic Skills Tests was a requirement for graduation in Minnesota at the time of this research. They are currently being phased out and replaced by a test that is used for both federal accountability (the reporting of adequate yearly progress) and for the receipt of a diploma in Minnesota.

many times they had taken each test, and seemed to understand the importance of passing. In addition, about 75% of students acknowledged using accommodations, with older students reporting doing so more than younger students. Two-thirds of the students were also able to identify other accommodations that would be helpful to them in the future, validating the "importance and need for research that addresses the perceptions and opinions of students who indeed face the greatest consequences as a result of participation in high stakes assessment" (Thompson et al. 2000, p. 15). In this example, special education students were given opportunities to understand the testing process and were able to articulate their experiences to researchers. Like so many studies inclusive of student perspective, though, the opportunity to engage students in a reflection of their experiences was short lived and extraneous to their daily life in schools.

In the second local study, focus groups with English language learners and with parents of English language learners from five different language groups were conducted in order to uncover knowledge of, experiences with and opinions of BSTs (Quest, Liu & Thurlow, 1997). The study was done in 1996, during the early stages of implementation of these statewide graduation tests. The authors note that their literature review turned up no resources specifically addressing what ELL students and parents thought about issues of high-stakes testing, making their study groundbreaking. Results of the focus groups showed that most parents and students favored testing, with participants believing that tests can show who needs more help. Participants also suggested different tests based on English proficiency level, offering prophetic advice given the requirements of No Child

Left Behind six years later. Students also commented on the overall difficulty of the reading test, and had suggestions for improvement such as avoiding idioms, using simpler vocabulary, including a glossary and pre-teaching words likely to be found on the tests. Other students felt the tests were not hard as long as they could take their time, and chastised their peers who didn't take the test seriously, though recognized that some "may have thought that they were not going to graduate anyway, or that the tests were too difficult, when they probably could have passed if they had tried" (p. 7). Most students were not aware that special testing conditions were possible for ELLs, and felt that students should take the test in the same way as other students. When other arrangements are available, however, students felt strongly that students should have a choice in the matter, rather than having a teacher or counselor decide on their behalf. Participants also agreed that an exemption of at least 3 years was reasonable for students who had recently arrived in the US. Finally, both students and parents appreciated the chance to talk about the BSTs and have some of their questions answered.

Each of the previous studies offers another glimpse into the experiences and perspectives of students taking tests and some of their reactions to those experiences. The scope of the research is limited in the exploration of how students had appropriated the policy. In the first, special education students were asked to reflect back on their opportunities for accommodated learning in class and on statewide assessments. The interviews were off campus, and followed a scripted format. In the second study, English language learners had the opportunity to express their opinions of the BSTs, and while

55

their responses were informative and valuable, individual experiences were not explored in depth. In each of the studies, outside researchers gathered the data, and the methodology prohibited an in depth exploration of individual factors; but each study did bring students to the table as informed stakeholders. To get at true sense making of testing policy, a deeper exploration into these issues is warranted, requiring different and more extensive methods of research.

Voice Studies

When given the opportunity for extended and collaborative reflection, students reveal a host of other phenomena associated with testing. Reflecting more deeply on the broader implications of testing, Susan Bernstein (2004) describes the experiences of college freshmen who were given the opportunity to opine about their high school testing experiences in Texas, and the "shape shifting" required of her students as they [made] "the transitions from standards-based education to college basic writing"(p. 5). Using "education" as the generative theme for a developmental writing course, Hispanic students (generation 1.5) were given the time to critically examine their education and the systemic variable that shaped their experiences. They began to identify the fundamental differences between their Texas high school experiences, which focused considerably on passing state-mandated standardized tests, and the expectations for intellectual inquiry demanded of them in college. In a case study of one struggling student –Noah—Bernstein (2004) shares some observations that illuminate the depths of the dissonance between

56

standardized k-12 curricula and collegiate expectations. In one particularly astute observation about his test-prep in school Noah states: "Public schools give us the TAAS test, which seems to include material that I think is not beneficial for college [sic]" (p. 18). Bernstein describes this student's "growing awareness of how his reading and writing were shaped by standardized testing" (p. 18), including the ways in which history was reduced to short reading passages which didn't allow room for the exploration of multiple perspectives. Noah "discovered that in his first year of college that his professors placed more value on critical thinking, analytic reading, and persuasive writing than on "finding the right answer" (p. 18), and suggested that students needed a more rigorous, college preparatory curriculum in high school. To overcome an inadequate high school education, students must be ultra determined and develop a clear sense of purpose. In describing her focal student and his need to repeat his college writing course, Bernstein writes, or; "Rather than conclude in despair, Noah grounded his reflections in advocacy for future generations of students in his community" (p. 19).

Given the opportunity to critically think and reflect about their experiences, the students in Bernstein's class developed a sense of agency after realizing their collective experiences were part of a larger system of education that had significantly shortchanged them. Becoming incensed at their missed opportunities spurred the development of skills needed to succeed in college. Sadly, students in the same class related stories of friends and relatives who had dropped out of high school, in part because "continued test failure" precluded them from many opportunities for a chance at academic success.

57

The discrepancy between what is expected of students in high school versus expectations in college has been duly noted by students still in high school, as well. In recent research about No Child Left Behind the voices of English language learners, teachers and administrators in ten New York City high schools are shared (Menken, 2006). Like Bernstein (2004), Menken encountered students who recognized the discrepancy between their high school education and what will be expected of them in college. As one student from an ELL magnet school noted, "We not really learning anything, we're only learning the content of the tests and not what we're supposed to know and go to college. So now this Regents thing is making our classes be prep classes. Not like real classes I'm supposed to be taking, so I feel very bad staying in class" (p. 536). Another student commented on the lack of enjoyment for learning that can be found in the prep courses, and the speed at which topics have to be covered. Overall, Menken concludes that curriculum and instruction have been tremendously impacted by the Regents Exams – especially the iterations post NCLB - resulting in an implicit change in language policy in each of the schools of her study. Test-prep is par for the course for English language learners, in both ELL and bilingual classes. In bilingual classes, if students are going to be tested in their first language, content is delivered solely in their first language with no balanced emphasis on English language development (these students still must pass the English language arts Regent Exam). In preparation for the English Regents Exam, ELL classes look very much like English language arts classes

58

for native speakers, leaving little room for content-based instruction or the development of communicative competence by newcomers.

In the previous two studies, students exhibited a keen awareness of the greater systemic impact of testing policy on their educational experiences. Their roles, however, remained rather passive. The next study reviewed offers a description of policy appropriation that includes a consideration of how students are *acquired* by policy, and the students' sometimes active resistance to this acquisition. Jill Koyama's (2004) study discovered ways in which high school English language learners made very deliberate attempts to assert control over testing policy in their own, defiant way. Using interviews and participant observation, Koyama explored:

> …how teachers, within the restricting organization of the school, take in or appropriate the federal, state, and district policy to delineate positions for ELL students to inhabit, as well as the ways in which students and teachers, together, negotiate and give meaning to these positions by selectively implementing chosen elements of the policy and excluding others. (p. 403)

She describes Mexican-descent students who refused to take the California English Language Development Test, some of whom walked out of the testing room without answering a single question as well as others who had strong misperceptions about the purpose of the test. One student thought the test was to show they were not "illegal [immigrants] or something", and others felt the test didn't really matter, yet they recognized that tests were part of the experience of school. Koyama shares reports of students claiming to score low on purpose in order to be placed in sheltered classes with their friends, even though this limits their access to classes with native English speaking

59

students. Viewed from a systemic lens, students' purposeful manipulations of the testing

process "ended up reifying their positions of "failure" constructed by the testing process

itself" (p. 410), though doing so also reified a sense of control from the point of view of

students. Quotes from a selection of other students amplify how students are "acquired"

by positions that have often been created for them by others. Koyama provides an

interesting angle from which to examine the process of policy "acquisition", setting up a

description of how policy, as it is interpreted first by teachers, takes in students only to

have students exert their own power, influencing the overall appropriation of the policy in

the creation of a tenuous, sometimes antagonistic relationship to the process.

As was demonstrated earlier, misunderstood purposes for assessment and

motivation play an important role in the how students interact with tests. Other studies

have shed light on the role of stakes; if the stakes are high for students (i.e. a diploma is

contingent on passing a test), motivation plays a different role than if the reputation of the

school is the primary stake. Students also have astute observations about changes in

curriculum as a result of testing, describing how their options have been limited and the

focus of the content narrowed. In some cases, however, students (including English

language learners) have taken matters into their own hands; actively choosing to

participate in the whole process, or defiantly taking a stand against tests – even

sabotaging the process.

CHAPTER 3

METHODOLOGY

The more researchers can appreciate students as thoughtful, inquisitive, caring people whose ideas and experiences they should seek, come to know, and take seriously, the more they can reconfigure ideas about who students are, what students can and should do, and what it means for students to be and to become productive and engaged in classrooms and schools.

(Thiessen, D. 2007, p. 8)

The authentic inclusion of students' voices in educational policy research poses a myriad of political and ethical implications, including issues of access and reporting. How does the research process, itself a complex social practice, facilitate or inhibit the voices of students, especially those of traditionally marginalized populations? The design of this study straddles all three orientations of student voice research as described in the review of the literature. To reiterate, those orientations are a.), a focus on making sense of experiences in school (in this case standardized testing), 2) the exploration of the topic from the point of view of "marginalized", "at risk", or otherwise "differently" labeled students and the critical examination of students who are not well served by schools, and 3) the inclusion of students as collaborative researchers, "where they both describe and evaluate reforms" (Thiessen, 2007, p. 39). By combining purposes and themes of more than one orientation "…researchers seek to better capture the dynamic complexities of students' lives in classrooms and schools" (Thiessen, 2007, p. 53).

61

In order to encapsulate each of these qualitative, student-focused orientations, case study methodology was used that incorporated methods that can be described as 'funnel shaped' in coverage (McCall-Perez, 1991) – meaning the focus moved from breadth towards depth as the study evolved. In the case of this research, that evolution was from participant observation (Phase I) to document review, inquiry groups and interviews (Phase II). The primary source of data emerged during the methods which sought depth, namely inquiry groups and interviews. Complete descriptions of each phase of the study follows.

Rationale

This research was designed as a case study because of the inherent value for "refining theory and suggesting complexities for further investigation, as well as helping to establish the limits of generalizability" (Stake, 2000, p. 448). Given the topic of this study, a case study was also chosen because it can be a "disciplined force in public policy setting and reflection on human experience" (Stake, 2000, p. 448), which this study does by examining the intersection of testing policy and student experiences. Additionally, case study methodology has the propensity to explore new areas of research when "not much has been written about the topic or the population being studied, and the researcher seeks to listen to participants and build an understanding based on their ideas" (Creswell, 2003, p. 30). The intent of this research is to do just that – to devote time to really listen

to high school aged English language learners and together with them, co-construct an understanding of statewide testing from the perspective of these key stakeholders.

The following design was created with the intent to allow for such a reciprocal environment, with an attention to language and the negotiation of meaning on the part of both researcher and participants throughout the research process.

Study Design

This case study is bounded by place (one suburban high school ELL program in Minnesota), policy (related to statewide testing), and time. The study took place from January to June of 2007 to coincide with the "testing window", or the time during which the focal tests were administered.

Site

The study took place at Aldrich High School[5], which is located in a suburban district bordering on a large metropolitan city in Minnesota. The research site was chosen because my degree of "local knowledge - mundane, expert understandings" (Yanno, 2000 as cited in Wright, 2002) could not only help gain the trust of students, but contribute to the overall analytical framework, as well. At the time of the study, I was in my thirteenth year of employment as an ELL teacher in the school district. I had taught for five

---

[5] All names of people and specific locations have been changed to maintain confidentiality.

consecutive years at Aldrich in the late 1990s. As a current volunteer in the school, I was able to observe the day-to-day happenings and develop a strong sense of the context. The following demographic information describes the site and other contextual factors.

District

The school district in which this study took place encompasses all or part of 7 different suburbs, bordering a major metropolitan area. There are 13,000 students enrolled in K-12, served by 11 elementary, 3 middle and 2 high schools as well as one alternative school. According to information reported by the district and made publicly available on the Minnesota Department of Education website, eleven percent of students are limited English proficient. The poverty rate is thirty-seven percent, based on the number of students eligible for free and reduced-price lunch.

School

This study was carried out at one of two comprehensive high schools in the school district. During the time of the study, all high school English language learners in the district attended Aldrich High School. There are 2,300 students in the school, of which 250 (11 percent) are English language learners, representing over 20 different language backgrounds. The overall poverty rate for this school is twenty-three percent.

ELL Program

The ELL program in this school is comprised of 7 differentiated levels of instruction. Two levels (named Literacy 1 and Literacy 2) are designed for students who speak a non-mainstream variety of English (Liberian, for example) but have interrupted

or limited former schooling. For those students who speak a language other than English there are 5 levels of instruction, with Level 1 being for newcomers who are at the beginning stages of language proficiency, and Level 5 for students who are mainstreamed in all other content areas. Additionally, there are a group of students referred to as "Transitional". These students are monitored for academic progress but are no longer registered for an ELL class. More often than not, transitional students are still attempting to pass one of the state's required graduation tests. There are 4.1 full time equivalent (FTE) teachers and 1 full-time educational assistant in the ELL department.

Participants

Criterion- sampling (Merriam, 1988; Patton, 2002) was utilized. Sampling criteria included:

- **English language learners,** as identified by the state of MN and therefore would have taken the Test of Emerging Academic English (TEAE), and been assessed on the Minnesota Student Oral Language Observation Matrix (MNSOLOM). These tests have been chosen by MN to meet accountability requirements for English language learners under NCLB.

- **Students in grades 10-12** who were still part of the Basic Skills Tests (BST) of reading, writing, and math graduation testing requirements, which are being phased out and replaced with new requirements beginning with 9th graders in 2007.

- **Students in Level 5** or deemed **"Transitional"** (as described in the previous section), the determination of which was made by the ELL department at the school based on a variety of measures. Level 5 is the highest level of ELL class offered, and Transitional students are those who are primarily mainstreamed yet continue to receive minimal support from the ELL department.

All students meeting the above criteria were invited to participate in the study via a letter from the school principal (Appendix A) and invitation (Appendix B) which included details of the study as well as the requisite consent forms. The invitation was mailed home, as well as distributed during an ELL class announcement of the research as part of the recruitment phase. A total of 78 invitations were mailed to the homes of students. Eleven students indicated an interest in participating and had completed the necessary consent/assent forms. A demographic breakdown of the participants can be seen in Table 3.1.

Table 3.1 Study Participants: Language Background and Years in U.S.

Participant	Language Background	Years in U.S.
Li	Chinese	5
Maria	Spanish (Mexico)	6
Xoua	Hmong	16
Genevieve	French (D.R. Congo)	3.5
Edgar	English (mother Liberian)	Born in U.S.
Arkadi	Russian	2.5
Belinda	Amharic	6
Winston	English (Liberian)	3
Sena	English (Liberian)	3
Jeanice	French/Swahili /Lingala (D.R. C.)	1 (studied Eng. In DRC)
Amina	Somali	6 (+ 3 years in Germany)

Tests in Minnesota

At the time of the data collection for this study, testing requirements in the state of

Minnesota were undergoing a change. The system being phased out called for the passing

of high-stakes tests of reading, writing, and math in order to receive a diploma. These

tests, the Minnesota Basic Skills Tests (BSTs), were first administered to students in 8th

grade (reading and math), and 10th grade (writing) if the students were present in MN

schools at that time. If not, students would then take the test at the earliest opportunity

available to them. The graduation tests allowed for an exemption of the passing

requirement for three years for English language learners new to the country. All English

language learners in MN also take the Test of Emerging Academic English (TEAE),

which consists of reading and writing components to test for language proficiency, a

provision of No Child Left Behind (NCLB). They are assessed for listening and speaking

according to the MN Oral Language Observation Matrix (MNSOLOM). Additionally,

depending on grade level, the Minnesota Comprehensive Assessments (MCAs) are given

to all students for accountability purposes under NCLB – reading in grade 10, and math

in grade 11.

For the sake of consistency, participants in grades 10-12 were invited to

participate because students in these grade levels fell under the diploma assessment

requirements listed above. Focus was given to the TEAE, MNSOLOM, and BSTs, as

these tests would have been taken by all participants. Current ninth graders were not

recruited, as they were under new and more stringent testing requirements for graduation. Accountability tests at the 10th (MCA- II GRAD reading) and 11th (MCA-II GRAD math) grade levels are now also the tests which will serve as part of the graduation requirements in the state. Figure 1 provides a graphic representation of the changing nature of testing requirements in Minnesota.

Figure 3.1 Changing Testing Requirements in Minnesota: MN GRAD Requirements, Following the 9th Grader

Student Grade	2006-2007	2007-2008	2008-2009	2009-2010
9th	GRAD Writing			
10th	Study Participants (BST Re-takes)	MCA GRAD Reading Grad Writing Retest		
11th		BST Re-takes	MCA GRAD Math Grad R/W Retests	
12th			BST -Retakes	Grad Retests

Adapted from: Minnesota Department of Education, Division of Research & Assessment (2007),http://www.education.state.mn.us/MDE/Accountability_Programs/Assessment_and_Testing/Assessments/GRAD_Component_MCA_II/General_Information/index.html

Data Collection

For this study, a "funnel shaped" approach to data collection was used (McCall-Perez, 1991), moving from breadth towards depth in light of the research questions. The use of the following multiple methods of data collection, including participant observation, an anonymous questionnaire (Phase I), document review, inquiry groups,

and interviews (Phase II) ensured an opportunity to view the phenomena under study from different perspectives.

Phase I

Observation of the testing process

By observing and volunteering on testing days, I was able to develop a frame of reference that could be used when talking with students about the testing process. Referents were established to use if necessary for prompting students (i.e.: the test you took in the gym…). This level of participant observation also assisted in laying groundwork for further phases of the study by building trust and a degree of familiarity with the researcher's presence.

Approximately eight hours over two days were spent observing and providing volunteer assistance during the MN BSTs of math and reading. These tests, originally given in the 8th grade and for which a passing score is required in order to receive a high school diploma in Minnesota, are often being retaken by students in high school, or taken for the first time if the student is new to the state.

Approximately 80 students (a combination of ELL and native English speaking students) were taking or retaking the BST reading test in one of the school's gymnasiums. In addition to observing, my volunteer responsibilities included helping with seat assignments, escorting students to the bathroom, recording test booklet numbers once testing had commenced, and distributing pencils. During the math BST, I was

69

assigned to a language laboratory in the school where 57 ELLs were taking or retaking the math test with the aid of an audio accommodation in English, Spanish, Somali or Hmong. My task on this day evolved from passing out test booklets and assisting with the technology, to eventually proctoring the exam in one half of the room for several hours.

Field notes were recorded during and immediately after the observations. These opportunities for participating as a volunteer provided valuable context and insights that aided in both the collection of data in Phase II, as well as in the analyses of all data sources.

Questionnaire

A brief, voluntary, and anonymous questionnaire was given to over 50 students (both ELL and native English speakers) after taking the test in the venues in which I was assisting (Appendix C) to ascertain a degree of basic knowledge of the test from a wide range of students in the school. Collected during the early phase of the study, these questionnaires added to the breadth of information obtained, and provided additional context and background information regarding testing in the school, generally.

Phase II

Document review

The transcripts of students participating in the inquiry groups were reviewed for statewide testing data as well as for overall school performance. Data included test dates and scores for all statewide testing, along with grade history and current class schedules.

Document review was done in consultation with participants during Phase II of the data collection, and used to verify and/or help clarify discussions regarding the number of tests taken by participants and when. Test records facilitated in the analyses of data, as well. Namely, they served as a reminder of the complexity of the testing system and the diversity of individual experiences with tests in terms of sheer numbers of attempts and ranges of scores.

Inquiry Groups

Jones and Yonezawa (2002) describe inquiry groups as "deliberate spaces in which students can analyze, reflect, and deconstruct the meaning of school structures and cultures" (p. 245). While the term 'inquiry group' is unique in research methodology literature, I resonated with the authors' description, especially in light of the sense making and student voice framework within which this study is situated. Rather than focus groups, which have a scripted question and answer format, inquiry based groups are "not about answering a question; it is a sense-making process about an issue, problem, or experience…. they are semi-structured spaces for authentic dialogue about lived experiences in schools and classrooms" (p. 247). Students have a sense of ownership in this type of dialogue; it is open-ended, and participants are invited to contribute questions, allowing better access to what Seashore Louis, Fabey and Schroeder (2005) refer to as "moderating variables", those multilayered, complex factors that influence the ways in which individuals make sense of a phenomenon. An inquiry group can provide the loosely constructed space "in which students can analyze, reflect, and

deconstruct the meaning of school structures and cultures" (Jones & Yonezawa, 2002, p. 245). This type of forum parallels Lincoln's (1995) support of the Socratic method, especially when working with students who have learned silence, namely it is "a form of questioning that develops, on the part of respondents, increasingly sophisticated, reflective, evaluative, and critical answers as the dialogue proceeds" (p. 89). Perhaps most importantly, inquiry groups allowed for the creation of a comfortable space in which participants' questions were welcomed and encouraged and conversations could flow, recognizing research as "a social process not only as an intellectual one" (Feuerverger, G., & Richards, E., 2007, p. 558). Rudduck (2007) refers to this process as 'consultation':

> …A form of student voice that is purposeful, is undertaken in some kind of partnership with teachers, and usually initiated by teachers. It involves talking with students about things that matter in school or gathering their views through writing…[it] may involve: conversations about teaching and learning and the conditions of learning; seeking advice from students about possible new initiatives…and inviting evaluative comment on school policy or classroom practice. Consultation is a way of hearing what young people think within a framework of collaborative commitment to school reform. (p. 590)

In order to facilitate such consultation, inquiry groups were conducted with 11 students in total, with no more than 4 students per individual group. Groups were assigned based on students' study hall periods, or on rare occasion, with prior approval from a student's teacher. Passes were distributed to students via their advisory teacher, and on occasion, in person by the researcher on the day prior to a scheduled meeting. For the Inquiry Groups, students were classified by the school as level 5 (English proficiency)

72

or Transitional, thereby guaranteeing a greater facility for participation in extended dialogue in English.

Each group met three times over the course of a four week period. Each meeting lasted approximately 45 minutes and was digitally audio recorded for later transcription. Light refreshments were provided for each group, and an inducement of a $10.00 gift card to a local department store was given to each participant at the end of the data collection. To the extent possible, the groups maintained some degree of stability in terms of participants. The nature of adolescent's lives and school scheduling, however, necessitated a degree of flexibility in terms of how the discussions emerged. Ultimately, participants were included in discussions for approximately the same amount of time, though the actual configurations of the groups shifted from time to time (See Table 3.2).

Table 3.2 Participant Attendance at Inquiry Groups

	Class Period 3	Class Period 4	Class period 5	Class period 7
IG1 4/17/07	Li, Winston, Amina, Jeanice		Belinda, Arkadi, Edgar	Genevieve, Xoua, Maria
IG2 5/9/07	Winston, Amina, Jeanice		Sena, Arkadi, Edgar	Genevieve
IG3 5/10/07	Winston, Amina, Jeanice	Maria[6]	Belinda, Arkadi, Sena	Xoua & Genevieve

Anticipatory tasks. Each inquiry group began with a simple written task (Appendices D-F), which served dual purposes. On the one hand the tasks served as an

---

[6] Maria often left school during her free period in order to work. She was able to meet with me during one of her class periods for an extended interview on 5/10.

anticipatory activity, creating a concrete way to focus students on the topic of the discussion and giving them some time to reflect on their own as they arrived before sharing with the group. While the tasks were focused on testing, they were not intended to narrow the focus of the subsequent discussion, though they served to provide a tangible reference from which the rest of the discussion could evolve. Having tangible references to which participants could refer also served as an ice breaker of sorts for each inquiry group.

The tasks were created via my teacher's lens, with recognition of my role as a language teacher during the process of research. The fist was created with specific deference to my first research question in order to get at students' knowledge prior to the inquiry group discussions ensued. The second task was a deliberate attempt to play with language a bit, while further exploring students' perceptions of the tests they had taken. Students were asked to complete sentences and create analogies. The third task was modeled after a project idea proposed by Roy (TESOL, 2007), in which students participate in the writing of multiple choice items in order to better understand the logic behind such tests. Overall, the tasks proved instrumental during data collection in setting the stage for conversation around tangible topics. These individual reflections were collected after each group discussion.

Interviews

Individual interviews with each of the 11 inquiry group participants offered a comprehensive picture of their previous schooling and other demographic information. This information is not often easily shared in a group setting, yet proved valuable in the interpretation of their stories. Examples of information gathered include languages spoken, extent of previous schooling, date of arrival in the U.S., and time spent in other countries or refugee camps.

All inquiry groups and interviews were transcribed, with particular attention given to the authenticity of voice for these English language learners. Given that "students are the experts on their own perceptions and experiences as learners" (Oldfather, 1995, p. 131), and therefore the "only authentic chroniclers of their own experience" (Delpit, 1988, as cited in Odlfather, 1995), their voices need to be represented authentically. Striving for authenticity in the representation of voice for immigrant students also brings issues of language and language register to the surface. 'Student' speak translated into adult words does not always have the same meaning, creating challenges for communication (Cook-Sather, 2002; Mitra, 2004; Schultz, 1997). Leung, Harris & Rampton (2004) contend that representation really begins with transcription, which does not attend to the prosodic features of language, reducing it to orthographic features filtered through the lens of the transcriber. Orthographic features can, however, make strides towards representing authentic voice, and care was taken in the transcription of the

data to retain the authenticity of voice. Though the initial transcription was hired out, I carefully reviewed each audio file, and edited the transcripts to reflect students' voices phonetically and to adhere to colloquialisms.

<div align="center">**Analysis**</div>

Using primarily field notes, transcribed audio files, and participants' school records, the analysis of all data was ongoing, and began as data were collected. Emerging themes were identified by the researcher and shared with participants during the inquiry groups, and continuously checked against the audio files as the analysis continued. Transcripts were inductively coded by the researcher: with recurring themes or common threads sought (Miles & Huberman, 1994). Upon completion of the data collection, all data were coded according to research question (Table 3.3). During this first, comprehensive sorting of data, it became clear that data revealed distinctive differences based on the tests they described. Therefore, the second iteration of coding focused on themes by focal tests (either the TEAE or BST). During this round of coding, some common themes emerged across tests in addition to some that were quite distinct. Finally an adaptation of the sense-making framework described in the literature review facilitated the third iteration of analysis. This warranted a global review of all data, driven by an underlying attention to research question three (how ELL make sense of tests). The analysis involved a review of the data in light of those reported characteristics that were individual in nature, those that were contextually influenced, and those based on signals sent by the policy and /or test themselves. This stage of analysis also prompted an

exploration of how the first two research questions are unique from, yet also subsumed

by the third question.

Table 3.3 Three Iterations of Analysis

First Iteration: Data by research question		
ELL knowledge of statewide tests (RQ 1)	ELL Experience with statewide tests (RQ2)	How do ELL make sense of tests? (RQ3)
⇩	⇩	⇩
Second Iteration: Themes by test (perceived stakes)		
TEAE NCLB Features of tests Purpose Consequence Scores	TEAE Perceived Rigor Testing procedure Lost instruction	TEAE Stigma Proficiency Level
		BST Perceived Inequities
BST Features of test Exemption	BST Test procedure Accommodations Lost instruction Perceived Rigor Multiple retakes	Generally Formative assessment Knowledge brokers Test fatigue
⇧	⇧	⇧
Third iteration: Application of sense making framework		
Individual Characteristics: time in US, language proficiency, number of tests taken, prior schooling		
Contextual characteristics: peers, teachers, research process, school environment, family		
Policy signals: Test specifications, test administration, accommodations, teachers as filters		

Perhaps the analysis can best be described as three-dimensional, and is best represented

in figure, versus table format. (Figure 3.2). The sense students make of testing was

explored through data by tests, as well as by research question.

Figure 3.2 Integrated Nature of Analysis

Sense Making
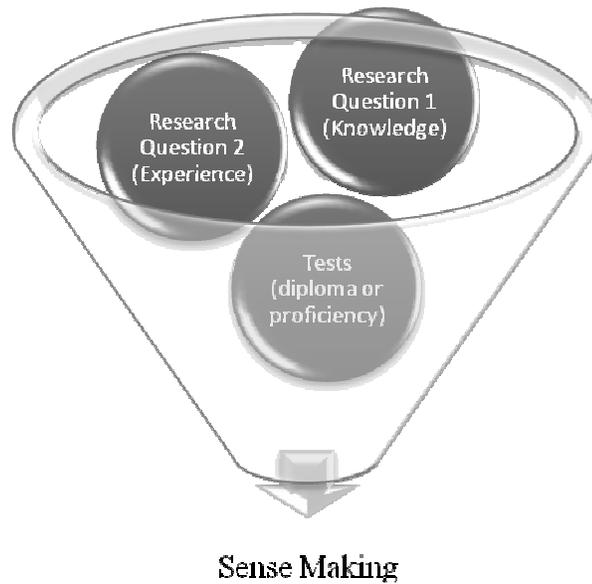
For each research question, inquiry groups were the primary data sources, though
other sources provided substantiating evidence. Table 3.4 provides an overview of data
sources by research question.

Table 3.4 Data Sources by Research Question

Research Question	Data Sources
What knowledge do high-school aged ELLs in a suburban Minnesota school district have of statewide assessments?	Inquiry groups, anticipatory tasks, interviews
How do high-school aged ELLs in a suburban Minnesota school district describe their experiences with statewide accountability testing?	Inquiry groups, anticipatory tasks, interviews, field notes from observations, student records
How do English language learners make sense of statewide testing?	Inquiry groups, field notes, interviews

Triangulation occurred by data source. Evidence of themes was supported across inquiry groups and over time within the same group. These themes were then substantiated by data from anticipatory tasks, interviews and field notes. Multiple data sources facilitated a holistic understanding of the data. Findings were then substantiated with quotes from participants to maintain an authenticity of voice (Anderson, 2003; Creswell, 1998). Following Yonezawa and Jones (2007) as a model, representative quotes were used to substantiate thematic representations of students' voices in order to elaborate on patterns among the voices.

Attention was given to the representation of participants in the findings. As Lee and Simon-Maeda (2006) note, "In a postmodern world, realities and subjectivities are multiple and continually in flux, and unless our methodologies reflect and adapt to these complexities, we will end up with reductive interpretations of our participants' circumstances" (p. 579). Cook-Sather (2002) refers to the need for researchers to represent student voices in the "fairest ways possible and, at the same time, avoid relinquishing [their] roles in interpreting findings" (p. 8). Inquiry groups, by definition, served the purpose of creating a space for the interpretation of information and the clarification of understanding among all participants. It is the researcher, though, who has a responsibility to the final product that ultimately cannot be ignored. In a case study, for example, it is the researcher who ultimately decides what is needed to understand the case, creating the "dressing of the case's own story" (Stake, 2000, p. 441).

To that end, great care was given to the analysis of data and representation of participants, as will be seen in the following chapters. It should be noted however, that even with great attention to fair representation in the final representation of data in this manuscript, I am still interpreting participants' own interpretations of events.

Researcher Identity

Researcher subjectivity throughout the entire process is crucial, requiring honesty to self and transparency to others. Peshkin (1988) describes subjectivity as an "amalgam of the persuasions" that stem from circumstances of class, statuses and values as they interact with that being investigated. Banks (1998) refers to subjectivity as the "biographical journeys of researchers" that influence the entire process of research. One's own journey, then, impacts one's agenda as well an analytical framework, which influences the way in which the research is designed from the beginning.

My experiences as a teacher of secondary ELLs have aided in my facility to gain access to and the trust of the students included in my research. This same experience along with the length of time affiliated with the same school district (12 years at the inception of the current study), resulted in a level of trust and access perhaps unavailable to someone from the "outside". In addition to the pilot study described below, I have conducted two other research studies at this site in which secondary ELLs were asked to share their opinions via interviews. One study examined how Liberian high school students in suburban ESL classes perceive their educational experiences. Another study involved interviewing older (18+) African immigrants regarding their experiences with

Minnesota's previous, dual-pronged approach to statewide assessment which involved traditional standardized tests of basic standards in reading, writing, and math, as well as performance assessments of high standards.

Yet despite having some degree of "insider" status with the school and the district, I was not a current teacher in the school, and therefore had no relationship with any of the participants prior to the study. At the start of the study, I was, in fact, an "outsider". Whether such a consideration was real or perceived, having such status did not release me from other traditionally "insider" responsibilities - especially given my role as a teacher of English. Several researchers have noted the inherent opportunities for language learning in the process of research itself (Cook-Sather, 2002; Fielding, 2001; Jones & Yonezawa, 2002; Leung et al., 2004). The act of research is an opportunity for the negotiation of meaning, and such active engagement with language contributes to the process of language acquisition and the development of metacognitive skills. In other words, the act of research blended seamlessly with the act of teaching, and was facilitated by my status in the school and the extent to which I was able to create a sense of trust among participants and with me. This trust was established through careful attention to the stages leading up to the research itself. During the recruitment phase, I had a colleague (who was an ELL teacher at the school at the time) introduce me to her students and facilitate a question and answer session. Additionally, at the beginning of each inquiry group time was allotted for informal conversation as well as for students to ask questions of me. These conversations – which covered topics ranging from prom,

pending graduation, Advanced Placement testing, and future aspirations - allowed for natural follow up at subsequent inquiry groups, and helped to create a sense of trust as a result of my genuine interest in participants' lives.

Trust can be an elusive concept when conducting research, requiring careful consideration of the degree of involvement one wishes to assume when researching with adolescents. In order to establish trust, issues of power and researcher identity had to be carefully considered and addressed. Encounters are always framed by power, and are especially significant to those who have less of it (Fielding, 2001). As noted by Graue and Walsh (1998), "…the boundaries between children and adults are much more elastic than many researchers seem to think (p. 80)". Boundaries need to be recognized, however, as there will always be a limit to the degree to which we can access someone else's world. My years of experience teaching adolescent ELL fueled my understanding of these limits, but also aided in my ability to attempt to reduce the limits while also recognizing them in my own research.

As a White teacher and researcher of culturally and linguistically diverse participants, the work of Schultz (1997) resonated with me. She believes there is "merit in acting as a cultural broker" between urban youth and a White teacher like herself. Adding that if the purpose of research is to bring about change, "…it is critical for researchers to cross boundaries to try to understand and write about people different from themselves as well as those who are similar" (p. 507). A careful consideration of the power of researchers in shaping the research process considers not only adult agendas

82

and frameworks for analysis, but the power adults have to create spaces where students voices can emerge, or not. The goal of the inquiry groups was to create those spaces.

Haw (1996) notes that in research there will always be a power differential, regardless of who is "doing the research and who is participating in the research" (p. 5). She emphasizes the importance of explicitly placing her subjectivity as a researcher in her work, along with critically exploring its limitations. She writes, "a recognition that we all speak from a particular standpoint, out of a particular experience, a particular history, a particular culture…does not then imply that we can only research the familiar" (p. 7). An awareness of my own subjectivity played an important role in the analysis of data and discussion of findings, as is evident the next chapters of this dissertation.

Pilot Study

In the spring of 2006, a collaborative pilot study was conducted at the same site in which the current study took place (Anderson & Stone, 2007). In the pilot study, exit interviews with English language learners following state-wide accountability tests were conducted, revealing the degree to which ELLs had basic knowledge of the intent and consequences of statewide accountability tests based on the following research questions:

1. What **knowledge** do secondary ELL in a suburban Minnesota school district have of statewide tests (MCAs, TEAE, BSTs)?
2. What do secondary ELL in a suburban Minnesota school district perceive as the **purposes** of statewide accountability tests?

3. What do secondary ELL in a suburban Minnesota school district perceive as the **consequences** of statewide accountability tests?

4. How do secondary ELL in a suburban Minnesota school district describe their **experiences** with statewide accountability testing?

The study provided both substantive and procedural points on which to build for the current study. While the topic of the current study remains ELL and statewide accountability tests, the focus (as noted earlier) has shifted towards that of sense-making on the part of the participants, with careful attempts to move from soliciting students' perspectives, towards encompassing all three orientations of student voice research (Thiessen, 2007). To do so, a less limiting design was warranted. Rather than one-time scripted interviews or focus groups with participants as the sole source of data (as was the case in the pilot study), a broader, more complex design was developed in order to create spaces for authentic dialogue and reciprocal learning as part of the research process.

The pilot study elucidated the varying degree to which English language learners in one setting understood the purposes and consequences of statewide accountability testing, and revealed a number of their concerns regarding process and content of the tests. These findings led to the current exploration of student agency in statewide testing and the sense they make of testing as active participants in the accountability process.

CHAPTER 4

FINDINGS

I think this is the first time a teacher asked me about the tests. It's good thing like to hear opinion from the students and then they can improve the tests.
Li (study participant)

An examination of the sense made of testing policy among all actors is warranted before its implications can be fully explored. This dissertation research is focused solely on students in the implementation of testing policy, specifically the sense they make of statewide testing in Minnesota based on their knowledge and experience with all aspects of testing. The knowledge and experiences shared by participants are on the one hand uniquely personal, and on the other, socially situated and influenced.

As noted in chapter 2, James Spillane and his colleagues (2002) have established a three-tiered framework of sense-making encompassing individual cognition, situated cognition, and policy stimuli as distinct yet equally valid facets of sense-making. All three tiers are considered in the analysis of data[7]. However, while the work of Spillane et al., does not focus on students, a nod to students is given in their description of the 'interactive web' of policy implementation. This web is complex and multi-dimensional, as is the process of sense-making. How actors in a policy's implementation (including

---

[7] For the context of this study, I have interpreted policy stimuli as those messages regarding tests' purposes and procedures as they are filtered primarily through teachers in the school.

85

students) assign meaning to the policy depends greatly on their "repertoire of existing knowledge and experience" (Spillane et al., 2002, p. 393). In this study, participants' *pre-existing* knowledge and experiences with testing are explored alongside the *new* knowledge gained as participants in this research study. In some instances this new knowledge came from the researcher, though more often participants co-constructed understandings based on their collective experiences and knowledge while sharing with each other.

This multi-dimensional nature of sense-making is woven into the presentation of themes in this chapter. Participants' basic knowledge of the focal tests (research question 1), their experiences with each of the tests (research question 2), and finally the overall sense they make of tests (question 3) provides the organization for these findings, with salient themes highlighted and presented for each research question.

Though the data are much less linear, the presentation of findings related to research questions one (knowledge of tests) and two (experience with tests) is organized by test. Early in the process of analyzing data, it became clear that the stakes connected to testing influence the ways in which students make sense of testing. For whom are the stakes highest? If for the school/system rather than for the individual student, participants' knowledge of and experiences with the tests varied. Thus, data from this research pertaining to knowledge and experience (research questions one and two) will be discussed first as they relate to tests required under NCLB (for school systems accountability), followed by data related to tests required by the state of Minnesota for a

86

high school diploma. Finally, the sense students make of statewide accountability testing in general (research question three) will be presented according to findings from each test.

Language Tests for System Accountability (TEAE/ MN SOLOM)

In this section, I address participants' knowledge (research question one)[8] related to the TEAE/MNSOLOM, followed by their experiences with the same tests (research question two)[9]. Findings are based primarily on inquiry groups and interviews, with supporting evidence from anticipatory tasks.

Basic Knowledge of NCLB

In Minnesota, the tests used to demonstrate growth in language proficiency for English language learners are the Test of Emerging Academic English (TEAE) and the Minnesota Student Oral Language Observation Matrix (MNSOLOM). This testing requirement is one of several with high stakes consequences to schools under No Child Left Behind (NCLB). I asked each group about the law to set the stage for subsequent discussions. As a federal law with great impact on the lives of students in school, NCLB appears to be of little immediate concern to participants. While none had heard of NCLB as an acronym, when spelled out there was a general recognition of the policy and its

---

[8] RQ1: What knowledge do high-school aged ELLs in a suburban Minnesota school district have of statewide assessments?

[9] How do high-school aged ELLs in a suburban Minnesota school district describe their experiences with statewide accountability testing?

overall intentions, though little understanding of the details. Participants in each of the

inquiry groups alluded to the basic premise of the law, as well as a few key tenets:

> Jeanice: I think they mean that every child has to go to school, yeah. Kind of like
>
> this.
>
> K (Karla): So to make sure everybody's in school?
>
> Jeanice: Yeah, and doing well.
>
> K: …and how do they know that children are doing well?
>
> Amina: Tests….they take the money away from schools. (IG2_P3) [10]
>
> Sena: Oh, I think it means no one is left behind, graduation, and, tests, and stuff
>
> like that. (IG3_P5)
>
> Genevieve: It's a program where, pushes student to go to school and stuff, I don't
>
> know. (IG3_P7)

Though each group was able to touch on some features of the law, details of its overall

complexity or even those details related to testing were unarticulated or unclear, creating

an opportunity for me to clarify. One participant, after expressing her belief that the law

was not working, elaborated by citing an example a teacher had shared with her of a

cousin who was very talented in the construction trade, but not good at taking tests, yet he

"might be left behind, but he can do so much" (IG2_P3). In sharing this example, Amina

not only employs the rhetoric of the law, but reveals input from a teacher in helping her

---

[10] The notation system indicates the data source first, followed by the class period during which students
participated. For example, IG3_P5 = inquiry group 3, class period 5.

reach some understanding. No other participants could recall any direct explanation of the law from teachers.

<div align="center">***Basic Knowledge of Language Proficiency Tests***</div>

Basic Features

With regard to tests required under No Child Left Behind to demonstrate gains in English language proficiency (in the case of this research, the TEAE and MN SOLOM), students had non-specific knowledge of the tests, at best. Despite the fact that teachers, administrators and policy makers in Minnesota tend to refer to the TEAE test by its acronym, participants in the study had no knowledge of what the letters stood for. A representative indication of confusion was given by one student:

Amina: Technology? Education? Academy? I have no clue. (IG1_P3)

During one of the inquiry groups, a discussion ensued regarding whether or not participants had even taken the TEAE test.

Jeanice: No, this year, I am not taking it.

K: Oh you didn't take the TEAE test?

Jeanice: No, I don't remember

Li: We are supposed to take it

Amina: We took it in like second quarter

Jeanice: Oh did we?

Li: We did take it.

Jeanice: We are taking too many tests that I don't remember the semester in my

head.

…

Li: Yeah, it was a month ago.

Jeanice: It wasn't the BST?

Winston: No.

Amina: uh-uh

Li: It is not BST

Amina: BST was in January

Jeanice: Oh! I took the TEAE test. (IG1_P3)

This exchange provides an example of how participants in the inquiry groups often

assisted one another in coming to an understanding of an event that affected them all.

Basic knowledge of this test was generated during the inquiry groups, situating

participants' awareness of a prior testing event socially, and during the research study. A

further example of this emerged during a discussion of which modalities the TEAE test

assesses:

Amina: Was it writing and math, or reading and math?

Li: Reading, writing

?: Nuh, there's no reading,

Winston: It was reading and writing.

Everyone: Reading and writing. (IG1_P3)

After a short exchange, all participants in this group came to a consensus that the TEAE does in fact assess their knowledge of reading and writing in English.

As for basic knowledge of the MN SOLOM (Minnesota Oral Language Observation Matrix), not a single participant was aware of the test, nor did they realize they had all been assessed yearly according to this matrix of speaking and listening ability, though one participant remembered getting the scores:

Sena: I got this delivered to my house, but I couldn't remember doing it. (IG2_P5)

Jeanice: Do we take it in high school or college? (IG2_P3)

In the case of this assessment, scores are being assigned, sent home, and factored into equations to determine whether or not a school district is meeting annual measurable achievement objectives (AMAO)[11], yet the students producing those scores are unaware of their roles in the process. Because students had such little knowledge of the MN SOLOM, the remaining presentation of findings related to accountability testing will focus solely on the TEAE – the reading and writing assessments of language proficiency.

Purpose / Consequences

Participants' responses regarding the purpose and consequence of the TEAE test varied. In response to a question about why students take the TEAE, Jeanice's answer indicates a basic level of understanding of the test's intent:

---

[11] Annual Measurable Achievement Objectives are required for all districts with English language learners. Accountability reports include measures of English language progress and proficiency, as well as academic achievement on states' reading and math tests.

Jeanice: I think, to see how, what is, your evolution, your year, and what effort

did you make, from the beginning and now, what changes did you make.

(IG1_P3)

Genevieve reveals that her information, while mostly accurate, is directly influenced by a

teacher.

Genevieve: Well, I don't know, I think, well, according to Ms. L., she said, it was,

it was a test to see how you are doing within ELL and out of ELL, with your

English …(IG3_P7)

With Xoua's contribution, a sense of shared responsibility among teacher and

student is evident, though Genevieve suggests the responsibility lies with the

student.

Xoua: Is it to see how well the teacher, the ELL teacher teaches you? And how

much you learn?

Genevieve: Or how well you're doing in your English progress. (IG3_P7)

Arkadi and Maria, like Xoua, place the onus on teachers:

Arkadi: To see where is like the school, something like that?...How the school

teaches the students. (IG1_P5)

Maria: That way they gonna know if they teach good or they not.

(IG3_P4)

With these examples, students are dancing around or very near the intent of the test – to determine whether students are making gains in English proficiency. The central role of the school and/or teacher in that process is indicated as well, with some participants describing the importance of individual effort, with others placing the responsibility for student learning with teachers.

Misperceptions of the TEAE's intended purposes exist among participants as well, and while Arkadi's response (below) is inaccurate, his response is indicative of the pervasive rhetoric of testing and resulting consequences of test scores, either within school walls, or from other societal inputs.

> K: … what do you think the purpose is for the TEAE test? …
>
> Arkadi: I think I heard like something about money for the school, and after
>
> school activities, and the students do good so they can get more money. (IG1_P5)

Whether the message is delivered overtly or covertly, Arkadi is aware on some level of the sanctions and financial consequences of many tests, though in actuality that connection is not as direct as Arkadi has indicated. Knowledge of the consequences of this test is informed by what students have heard in schools, whether from teachers or other students. There is a general understanding that the consequences of this test are greater for schools than students. Maria alluded to the lack of immediate consequence to students, and how this impacts her interaction with the test.

93

> Maria: Well I don't know, it's um, for me I don't know why they have to take it,
>
> they say, it's not, we don't get even credit, we just, they have to see, even when I
>
> was taking this test I was just playing. (IG3_P7)

If not for credit, or some other direct consequence, Maria is clear about her effort level on the test, a sentiment documented by others researching large-scale testing (Chabran, 2003; Flores & Clark, 2003; Koyama, 2004; Magnuson, 2003). Xoua, too, laments the lack of a clear purpose for this test. In sharing verbally her response to an item on Anticipatory Task 2, she noted that when schools give tests they should,

> Xoua: Like, make sure the student know, know about the test and what is
>
> it for, cause for the TEAE test, we don't really know what it's, like why
>
> we taking it, and what's the purpose, like we don't even know what AE,
>
> like T-E-A-E (pause)
>
> K: Stands for…? (IG3_P7)

Maria's and Xoua's use of the collective 'we' in the previous examples indicate students do talk about tests among themselves. Li's response supports this, too.

> Li: And then, I also heard the opinion from other peoples, and then they
>
> thought they wasn't really care about the test, cause it doesn't really affect
>
> you anything, cause you just like, you just take it. (IG1_P3)

Students are talking about tests and, as Xoua notes, would like to know the purpose. At the time of the data collection for this study, Xoua – a senior – had taken the TEAE test three times without knowing why. As noted by Spillane et al., "There is a critical need to

94

structure learning opportunities so that stakeholders can construct an interpretation of the policy and its implications for their own behavior" (2002, p. 418). While the authors were referring to teachers and their need to construct interpretations of policy, these findings support others who have found that doing so with students is imperative for a policy's overall success (Roderick & Engel, 2001; Rudduck, 2007). Failure to do so is a clear indication that students are viewed as mere recipients of testing policy – the TEAE test is given to them without their active consent, cooperation or even understanding. Yet when the purpose is clearly explained students are apt to go along with the policy. If given agency, students are more likely to participate than simply acquiesce. This was evident in my discussion of the TEAE with Genevieve:

> Genevieve: Yeah. Well, if I knew that part of the reason I was gonna take the test this year, because actually helps, like it actually helps, the government to see that student who learn English as their second language or whatever, do really, like, make improvement, you know, so, I, if I knew that I was gonna take the test. Cause I actually kind of like the fact that it does show that. (IG2_P7)

Generally, participants had partial information about basic tenets of language proficiency testing policy and the corresponding tests. There was more doubt, than certainty, however – sentiments which carried over to discussion of the scores for this test.

Scores

A discussion of the test's purpose emerged again in light of the scores for the TEAE. The discussion of scores began with trying to recall if participants had even received scores for this test, and what those scores mean. Admittedly, there was confusion with regard to the interpretation of their scores.

K: Do you remember getting scores last year from the TEAE test?

Winston: I can't remember.

In this group, Winston had no recollection of ever getting his scores on the TEAE, whereas Jeanice and Li remembered scores, but expressed confusion over their interpretation:

Jeanice: Yes. I have one, but I don't want…(giggles), I tried to read that, but I just see level 5, level something, English level something, I say what? I don't understand, I don't get the results, how did they do that

…

Li: I think that you take the test to put you in the right level, I think the results she talking about 5 tell you the level you in, so…(IG1_P3)

In fact, a level 4 or 5 on the listening/speaking and a 5 on the writing portion of the TEAE indicate a student is proficient according to the Minnesota English Language Proficiency standards. (For the reading portion of the TEAE, a score of 4 is considered proficient.) There is not, however, a direct alignment between TEAE scores to the levels of ELL classes available to students at this site, as was thought by Li in her response to

96

the question. Sena – a transitional ELL in 11th grade – had a similar misperception about alignment of TEAE scores to ELL level, expressed in terms of her own worry of being reassigned to ELL classes.

> Sena: I was, I was kind of thinking, that if the score comes back and you get a low score they might put you back in ELL,
>
> K: Ok,
>
> Sena: Well, cause for me, I have regular English right now, I left the ELL when I was a freshman. (IG2_P5)

Previously in the dialogue, Sena had indicated she was mad when she heard about the TEAE. Her misconception of the consequences unnecessarily evoked anger and frustration and is an example of the testing process itself constructing a position of failure for this student (Koyama, 2004). This exchange led to an opportunity for me to clarify, thereby impacting Sena's knowledge of the purpose and consequences of the TEAE test, essentially allaying any fears she had of being "put back in ELL". Despite my clarification of the consequences of the TEAE (and the relatively low stakes for her personally), she indicated she would continue to give her best effort.

> Sena: In my mind I was going back, if I don't do good, I was going to be placed in ELL, so I do my best.
>
> K: Alright, what about now, knowing that, they're not going to pull you back to ELL,
>
> Sena: I will still do my best. (IG2_P5)

97

Sena and Genevive represent students who are likely to commit to a reform or policy if given adequate explanation (Rudduck, 2007), while other participants are more likely to remain swayed by the lack of personal consequence, despite teacher or other adult input (Koyama, 2004).

The delay in receiving scores was admittedly a factor in participants' recollection of them, as represented by the following exchange with Li. Her lack of concern for the scores is also evident, and reiterated by Jeanice:

Li: The score take about, like half year to send it to you. You kinda forgot it. You already changed.

K: Do you worry about the TEAE scores?

L: I don't. (Int_L)

Li: We can't even remember what test is it, it's like…As you look at it and like, what is the result about? When did I took the, take the test? I was like, what?

…

Jeanice: When they said it was the TEAE test I just said AWW, I don't really care. (IG1_P3)

Lack of clarity or concern for the consequences of the TEAE test was expressed by Genevieve, too. She felt the scores had little use for teachers.

Genevieve: One thing I also don't understand is that if you take that test, let's just say you take that test and you don't do good, ok, what's the, what's the teacher gonna do about it? Cause I've never seen, like, a way where a teacher try to do

98

something about it cause we never actually got our score, or we get it like way

way later, and we even forget about that we took the test, cause I remember, I got

my, what last year's TEAE test like a couple of month ago, probably,

K: They came pretty late, didn't they?

Genevieve: So, yeah, so it's like, now that they see how I did good or bad, what

are they going to do about it? (IG1_P7)

Genevieve sees the futility of ascribing any real meaning to the scores on this test. This

same futility was described by participants in Chabran's study (2003) based on similar

complaints of delays in scoring and reported lack of motivation on a test where the stakes

lie elsewhere. When tests are large-scale, or statewide in the case of the TEAE, the time

needed to score student writing renders the scores of little formative value. Not only does

Genevieve perceive little benefit of the scores for teachers, she goes on to note how there

is even less impact on students like herself, given that she is no longer receiving direct

ELL instruction, a distinction that emerged throughout the analysis of data.

Knowledge Affected by ELL Status

Current status as an ELL student (transitional or receiving direct service), as well as

length of time in the United States impacts the sense students make of testing.

After reviewing the scores and the descriptions (retrieved from the Minnesota

Department of Education website) of the indicators for each level, and asked whether

they felt the scores were an accurate reflection of their English abilities, responses varied.

Amina, who had been in the U.S. six years, felt they were not:

Amina: I can read better than that. (IG2_P3)

This mirrors the perception of tests as inaccurate reflections of students' motivation or ability shared by students elsewhere (Yonezawa & Jones, 2006). On the other hand, Jeanice, who had just arrived in the country the week prior to her first TEAE test and subsequently had a low score, had a different opinion. She said, "I agree because last year I just came and yeah, I agree with that" (IG2_P3).

In sum, participants had little concrete knowledge of the TEAE, and even less of the MNSOLOM. Differentiated status as an ELL student and the amount of time participants had been in the U.S. was a factor: one that emerged again as students described their experiences with the TEAE.

In the next section, findings are shared in relation to participants' experiences with the TEAE (research question two). Data primarily come from inquiry groups, with supporting evidence from interviews, observations, anticipatory tasks, and participants' test records.

Experience with the TEAE (Research question 2)

The T-E-A-E test make me feel stuiped when the teacher gave me the passed to go take the test. (AT1_Xoua)

This quote, taken from Xoua's written anticipatory task during the first inquiry group reveals the strong affective impact the TEAE can have on students. Xoua had been in the United States sixteen years at the time of the study, and therefore had received all of her schooling in US public schools. She had not been receiving direct ELL classes for

100

several years, yet still qualified as limited English proficient (LEP) primarily because of her failure to pass a Minnesota graduation testing requirement – the Basic Standards Tests (to be described in detail later in this chapter). She also had not yet scored proficient on the reading portion of the TEAE. As a result of her status as LEP, it was imperative (from the school's perspective) that she take the test in order for her to be counted among those students who generate extra funding from the state department of education.[12] As noted previously, Xoua was unaware of the test's purpose, and felt stigmatized by being required to take a test she later described as easy. In this next exchange, Xoua and Genevieve (both transitional ELL) describe their affective reactions to the TEAE.

> Xoua: Cause, when they take the student that's out of ELL back and take that test it make them feel like really low, like, why am I still taking this? Am I still like, it seemed like they're putting you in a group that you cannot, it made you feel different from the other student, because when they keep bringing you back to the like, the level.
>
> Genevieve: It's kind of like, the group that you actually got, worked hard to get out of, it's kind of like feels like they're bringing you back into something you knew that, oh yeah, I had to do that, but now I'm out of it and I need to move on, you know,

---

[12] One factor used to determine additional state funding for English language learners is a score below proficient on the TEAE test.

K: Ok,

Genevieve: So it feels like they were bringing you back to whatever you were

trying to escape. (IG1_P7)

Despite Xoua's and Genevieve's claims of the TEAE being easy, both had recently taken

it again because of failure to achieve a proficient score the year prior. Neither had passed

the reading. In fact Xoua received a score of '3' four years in a row (a '4' is needed to

pass). Having been in the US 16 years, she meets the definition of a long-term English

learner (Freeman & Freeman, 2003), whose tenure in school failed, on some level, to

account for her unique needs as a language learners (Olsen, 1997; Valdez, 2001). One

might question whether the bar is set exceptionally high for the TEAE, or if lack of

motivation resulting from feeling stigmatized leads to lower scores. Ironically, Xoua had

passed the diploma reading test, but not the TEAE reading. Conversely, she had passed

the TEAE writing but not the diploma writing test. In other words, on the test of language

proficiency, her writing skills exceeded the target, yet she could not pass the diploma

writing test. Is the bar for the TEAE writing set too low? One might also question if the

raters for statewide writing assessments have had adequate training – whether scoring all

ELL writing (the TEAE), or predominantly non-ELL writing (the BST).

Perceived Rigor

Other participants commented on the level of challenge or rigor (or lack thereof) of the test, which factored into their experience during the test itself. Their perceptions of rigor correlate with individual characteristics such as time in the U.S., language proficiency, and language background.

Previous schooling and a first language of English are important factors in Winston's interpretation of his experience. While relatively new to the US (3 years), he came from Liberia, spoke English, and had relatively consistent opportunities for schooling prior to his arrival in the US.[13] He describes how he felt while taking the TEAE:

> Winston: I was relaxed.
>
> K: ..OK, why were you relaxed?
>
> Winston: Because I knew most of the stuff and it was so easy. (IG1_P3)

While the test itself caused Winston little anxiety, he later described it as a "waste of time" (IG1_P3), as does Amina:

> Winston: I think it was easy. It wasn't hard.
>
> Amina: It was a waste of time, technically …
>
> K: Too easy?

---

[13] A large number of Liberian refugees reside in Minnesota. Many speak Liberian English as the primary language, but due to civil war and a collapsed infrastructure at home, many of the children have significant gaps in schooling and qualify for additional literacy instruction through ELL programs.

Amina: Everyone was done before the time limit…[E]veryone was like

napping…I napped!

K: You napped?

Amina: Long nap, it was pretty good. (IG1_P3)

As a native English speaker, Winston is an anomaly among participants, though his

experiences are mirrored by other participants. Amina, who had been in the U.S. six years

and in Germany for three years, also describes the TEAE as a 'waste of time'. In a later

inquiry group, Amina and Webster described how the lack of challenge and consequence

to them affected their engagement with the TEAE test.

K: …You did not try? Ok why not?

Amina: It wasn't challenging enough. I want to be challenged.

K: Ok, you want to be challenged. So if the test is not challenging enough, you're

not even...

Amina: I'm not even going to try.

K: ... you're not going to try. Ok. How about the rest of you, did you try?...Did

you take it really seriously?

Winston: Not really.

K: Not really?

Winston: Cause I knew the score wouldn't count, so,… (IG2_P3)

For these students, the test is perceived as being not only irrelevant, but unchallenging as

well, a sentiment echoed by Xoua (in the US 16 years):

Xoua: And, I think um, the part on like the test, when you look at the picture and write down what they tell you, like what's on the table and stuff, it just make you feel like really stupid, because that seems like first grade, like, would do that. What's on the table and you gotta write down like what you see in the picture.

K: So it's way below your ability. Your level.

Xoua: Yeah…Cause I had a lot of friends who came out of there saying that was the easiest test that I ever taken in my life. I don't even see the point of it.

(IG1_P7)

Upon further discussion, participants in one group noted the folly of starting everyone at the lowest level of the test, regardless of what their actual level of English proficiency is. Given the timed nature of this test, students at higher levels of English proficiency often finished sections very quickly, and were forced to wait until they were officially instructed in order to move to the next section. Thus, Amina's recollection of a "long… pretty good" nap (IG1_P3). Amina had a solution to this problem however:

Amina: They should write the test so like levels, like ELL 5, they should make a test for that one and ELL 4, and so on and so on.

K: Ok, so have a different level test based on your English level.

Jeanice: Yeah.

…

Li: It says uh, they make the test only for the ELL students, not English speakers, so I think they should put it like, level by level. (IG1_P3)

105

Li notes that since the test is designed only for ELL, a recognition of the proficiency

levels of students taking the test is warranted – a belief shared by students in other

research (Quest, Liu & Thurlow, 1997). She described the benefits of the TEAE for

students whose proficiency level is lower, and the positive affective response a test can

provoke if it is leveled such that students can demonstrate success.

> Li: Umm, I feel like, for the test, I can see it's like, it's like, they write it at like, 5
>
> levels, it's from easy to harder, like getting harder, but the harder one, it wasn't
>
> like really hard, so it would make me feel like I'm getting it, like some tests, it's
>
> really hard, and you feel like some pupil might feel so bad, like mostly for second
>
> uh, people who have second language, and then they thought it was like, kind of
>
> like, I'm dumb, why I don't know the language,
>
> K: Cause they are too hard?
>
> Li: Yeah, And then this test make me feel like I'm getting there, I know the
>
> English right now. (IG1_P3)

Li felt successful and confident in her ability to demonstrate her English skills, yet had

only taken the TEAE two times. While she had been in the US several years, she had

been in Minnesota just over one year. During one of the groups, she gave a description of

a similar test of language proficiency she had taken in another state. For Li, the TEAE

was an affirmation of her English proficiency abilities. This is in contrast to some of the

other participants who, having achieved relative success in language proficiency and

other academic indicators, felt the test was beneath them.

Genevieve commented on the lack of challenge of the TEAE for students out of direct ELL classes, like herself. In Genevieve's estimation, success in classes with native speakers is a strong indication of her overall academic language proficiency, a sentiment shared by Xoua:

Xoua: And I think the TEAE test should be made for the student that's still in ELL only, because if they out of ELL, they should like, based on their class, they should just judge the student by their class and their grade. (IG1_P7)

Genevieve extended this train of thought, noting that transitional ELL students often perform better academically than native English speakers:

Genevieve: …well, in some classes, there are, like native speaker who do not do good in certain classes, and you might as well be doing better than a lot of them. (IG1_P7)

Her overall feeling that success in mainstream classes is a better indicator is augmented by the revelation that she is in Advanced Placement (AP) science classes!

Genevieve: …I don't see the point that I have to take this test because I'm already out of ELL and I have AP classes, so wouldn't I be like, an example of how good I am doing, instead of having to take this test again? And again and again? (IG1_P7)

Examples like those shared by Winston, Xoua and Genevieve warrant a careful examination of the TEAE as Minnesota's only measure of language proficiency. Indicators such as those mentioned by these participants, including success in advanced

107

classes, English as a first language, and extensive time in the US and mainstream courses support a more flexible, differentiated approach to documenting language proficiency. Students such as these are included in the testing because they have either been designated as ELL within the past two years per flexibility in NCLB[14], or remain transitional ELL because of difficulty passing state graduation tests, despite their relative success in non-ELL academic classes.

Experiences with Testing Procedure/Protocol

The previous discussion highlights the importance of considering proficiency levels when testing students. Several participants also described another factor to consider when testing students of various proficiency levels in the same setting: peer pressure. Students who may be at a lower level of English, and therefore need more time to finish each section of the test, notice and are affected when their more proficient peers begin finishing a section. Amina described her competitive nature, and how this made her feel she needed to go faster if others were finishing, leading her to note, "It's not fair how they put us all together" (IG1_P3). Jeanice's reaction to the same phenomenon was: "I'm like wow, am I not smart or what?" (IG1_P3). Students are very alert to their surroundings during a test, and for students like Amina and Jeanice (who were both still in ELL classes), self-doubt emerged about their own academic abilities. As students hurry

---

[14] States are allowed to include "former ELL" for up to two years in Adequate Yearly Progress determinations. (Non-regulatory guidance, Office of Elementary and Secondary Education, US Department of Education, May 2007).

to finish, whether due to peer pressure or the time limit itself, how does this impact the validity of resulting scores? And what of the emotional impact on students when self doubt, and possibly stigma begin to set in? The previous section explored students' experiences with the TEAE, and how the proficiency level of students at the time of the test is a real factor in how the test is perceived.

Lost instruction time. The most salient theme related to testing procedure at this site has to do with the timing of the test's administration. English language learners taking the TEAE are released from class for the test while the rest of the school (~89% non-LEP) carries on as usual. In referring to her friends, Xoua shares:

> Xoua: And, a whole bunch of them said that, the only want to take it to
>
> skip a couple hours of classes.
>
> K: Oh, so they did it to get out of class.
>
> Xoua: Right, so basically that tells you all about the test. (IG1P7)

While Xoua's friends looked at the TEAE as an excuse to get out of class, participants in the study eluded to just the opposite phenomenon, noting how the TEAE forced them to miss out on other learning opportunities and to fall behind in their schoolwork.

> Li: it's OK if the teacher doesn't' teach any new stuff for that day, but if they do,
>
> you gonna miss a lot of thing. (IG1_P3)

Li describes her conflicting feelings of not wanting to miss class, but the need to do so in order to take the TEAE, adding she was the only person in her chemistry class who had to take the TEAE. With a low incidence of ELL in many advanced academic content classes

at this school, teachers of those classes are undoubtedly less attuned to the policy's

affecting English language learners. Genevieve describes her deliberate avoidance of the

TEAE in order to attend her science class.

> Genevieve: No, this year I didn't take it….Although Ms. K was running,
>
> K: She was, running?
>
> Genevieve: Yeah, she was trying to get (?)
>
> K: So you were trying, you were hiding from her?
>
> Genevieve: No I wasn't hiding, I told her that, well that wasn't my fault, because I
>
> was about to take it, but then I couldn't miss my Chemistry class. So I went to the
>
> lady and told, I went to Ms. (?) and told her that I couldn't take it that day, so I
>
> could take it the next day or my 7th hour, and she said it was ok, but then I, on my
>
> way down there to find her, I met Ms. L. who told me that I didn't need to take it,
>
> so, since I didn't want to take it anyways, it was good for me.
>
> K: So, you were fine with that.
>
> Genevieve: Yeah, but Ms K. came after me, for a day, and then she's like, you
>
> need to take it, but I never went and she forgot about it, so it's ok with me.
>
> (IG1_P7)

Genevieve's avoidance of the TEAE was less of a deliberate attempt to thwart the testing

policy (Koyama, 2004) as it was concern for her academic status in her science class. Her

decision to avoid the test was exacerbated by conflicting reports by teachers of her need

to do so in the first place. In the same inquiry group, Xoua and Maria reported having

taken the TEAE as directed, but expressed frustration at having to make up the work from classes missed:

> Xoua: … we missed like almost like 3 hours
>
> Maria: 4 hours.
>
> Xoua: 4 hours, of class, and we gotta make it up on the next day, and like, we just, so behind on the next day it just, (IG1_P7)

The additional burden of makeup work was shared by other groups as well.

> Jeanice: Yeah, people should be in class doing their math, like, if you just missed that 2 hours of math like that, you're gonna be in trouble.
>
> K: Ok, so you're missing out on a content class, like math, did you have to then catch up and do all of that as extra homework.
>
> Amina: That's the worst part of the TEAE test
>
> Jeanice: Catch up! That's hard. That's hard to catch up, it's not easy if you have homework and stuff.
>
> Amina: And you have stuff in other classes to worry about. (IG2_P5)

In the following exchange, the same girls continue to discuss the fact that some things can't be easily made up, leading to the suggestion that classes be suspended for the TEAE, or the test be given outside of school time so as not to miss out on any learning opportunities. Here we have dedicated students willing to takes tests outside of the school day, rather than miss what they consider to be valuable learning opportunities!

Li: Yeah, I remember, the day I was taking it and then I had a very important lab in my class, so I just miss it, and then it's hard to make up cause it's like...

K: … were you able to get caught up? Did you find a time to do it or something else?

Li: Umm, we have a lab report, and then I kind of get idea from my partners, whatever, but I think...

Jeanice: Yeah, I think, if they want people to take this test, they should find a time out of school time.

Amina: every class should just take a break, like in ELL, we stopped doing everything. (IG2_P5)

As seen from the previous exchange, criticism of the test often led to opinions being shared in the form of suggestions for improving testing of language proficiency. In fact, these participants echo what researchers have reported (Magnuson, 2003; Quest et al., 1997), or will eventually conclude after years of research and millions of dollars (Yonezawa & Jones, 2006).

In sum, participants see no purpose for the test if the stakes are not personal and immediate (Chabran, 2003; Flores & Clark, 2003; Koyama, 2004; Magnuson, 2003; Nichols & Berliner, 2007). They reiterate Magnuson's assertion that students "by and large can ignore" system accountability tests such as the TEAE (2003, p. 191), which in turn challenges the validity of the TEAE scores. They viewed their time in class to be far more valuable than time spent taking the TEAE, and clearly articulated their frustration

with taking a test that was not adapted to their proficiency level. As a result of such non-adaptive proficiency testing, participants who reported "playing", or not taking the test seriously end up, paradoxically, reinforcing their position as English language learners (Koyama, 2004). Additionally, participants whose proficiency was considered advanced felt stigmatized by a test that starts significantly below their level. Magnuson (2003) describes a similar phenomenon, though for students at the opposite end of the proficiency spectrum – English language learners with limited formal schooling, and their frustration with being tested beyond their level. Participants' suggestions for attending to such issues, and the overall sense students make of this test will be discussed later in this chapter under research question three. I now turn to data that pertain to tests of greater stakes for students.

Tests for Graduation

In this section, I will revisit research question one (what knowledge do ELL have of tests) as well as research question two (experiences with tests) in light of the Basic Skills Tests (BSTs). In contrast to the TEAE, there was far less ambiguity among participants when the perceived stakes of the tests were high, and very personal. In Minnesota, students must pass tests of reading, writing and math in order to receive a high school diploma. As a result, participants in this study have a vested interest in these tests, as is evident by the relative clarity of their knowledge of basic facets of the Basic

Skills Tests. Findings related to knowledge of the acronym, the purpose and consequences, and allowable exemptions are presented below, and based on data from inquiry groups and anticipatory tasks.

Basic Knowledge of Diploma Tests

Overall, relatively little discussion took place in inquiry groups around knowledge of the BSTs. As noted above, there was little ambiguity or need for clarifying discussions among participants in relation to the first research question. There was no doubt that a diploma was at stake. Amina's reaction is indicative of the straightforward, matter of fact nature of discussions of the tests' acronym:

K: What does BST stand for, first of all?

Amina: Basic skills test.

K: Basic skills test? Ok, and what types of tests are there, for the BSTs?

Amina: Math, reading, and writing.

K: Math, reading and writing. And so, why do you have to take that test?

Amina: To graduate.

K: Ok, to graduate.

Amina: It's sad. (IG1_P3)

Amina's response is also representative of the emotion with which participants' responses were laden. Her comment about the test being 'sad' was in reference to a friend who had completed credit requirements the year before, but still had not received a diploma because of at least one Basic Skills Test.

114

Knowledge of Exemption

In addition to the difference in perceived stakes, another distinguishing factor of BSTs from TEAE/MN SOLOM tests is the allowance of an exemption for student in the US fewer than three years.

> Jeanice: Yeah, I didn't really need to pass it, because of, I have less than three years,
>
> K:…are you all aware of that?
>
> Amina: Yeah, you're not here for 3 years ...
>
> Winston: I didn't know that.
>
> K: You didn't know that? Ok, so can you, can one of you explain what it, what the rule is?
>
> Jeanice: Hm, the rule is if you have less than 3 years in the US, you won't, you don't need to pass the BST.

After some verification of Winston's time in the U.S., it was determined that he would not have needed to pass the BSTs to graduate, though at the time of the inquiry groups he had passed all three tests. Jeanice learned of the exemption from her teacher, but wanted to pass the test anyway:

> Jeanice: Yeah. Uh, the teacher told me because I didn't pass the writing first of all, she was sad, she was like, oh, I know but like, try your best for this, maybe you don't have to pass the BST, to graduate, I say ok, I will try my best. And so me, for me, I wanted to pass it, yeah, I really motivated to pass it. (IG2_P3)

The difference in knowledge of this exemption among students in this group may be a reflection of their ELL status. Winston was no longer in ELL, whereas Jeanice was. ELL students are not only more likely to be aware of exemptions, but to engage students in conversations about them, as well. Students in the fifth period inquiry group were not aware of the exemption policy on BSTs, and though it no longer applied to any of them, they supported the policy:

> Sena: I think that's fair.
>
> K: Ok. Why is that fair?
>
> Sena: Because they, they came kind of late, you know, we've been here for a little while, maybe we, probably gotten used to it, the American way of teaching…
>
> K: Why do you think it's fair?
>
> Arkadi: Cause it's helps them, maybe it gives them more time to learn before they take any tests. (IG2_P5)

Genevieve was able to elaborate on what is, in fact, the basic rationale for having an exemption:

> Genevieve: Well just because you haven't been here for long, so basically, let's just say, you come in 10th grade and you have level 1, of ELL, so it's not going to do you any good taking the BST because basically you will not be able to understand half of the stuff you're reading to be able to answer the question, and pass the test as you will, especially reading part, or writing. (IG2_P7)

116

The state of Minnesota has recognized, minimally, that language is a factor in students' ability to demonstrate academic proficiency on graduation tests, a factor these participants could articulate based on first hand experience. However, the policy for exemption erroneously associates time in the U.S. with proficiency level, implying that English language learners acquire proficiency at the same rate – an assumption that contradicts research in the field of language acquisition (TESOL, 2005). While not all participants were aware of the policy for exemption, they were all acutely aware of other basic features of the diploma test, for which the stakes were extremely high and personal.

Whereas the overall presentation of findings regarding the TEAE had as much to do with negotiating understanding and building knowledge of the fundamental features of TEAE policy (rationale, purpose, scores, consequences), findings related to the BSTs had much more to do with experiences. In this next section, I revisit the second research question (ELL *experience* with statewide testing) in light of the BSTs. With the stakes of these tests perceived as extremely high for participants, reported experiences were far more individually nuanced. Experiences with testing procedure and accommodations are presented, followed by an exploration of the impact multiple retakes has had on students' experiences. Data sources include observations, inquiry groups, anticipatory tasks, interviews, and participants' records.

The data reveal what others purport to be true based on teachers', administrators' and researchers' accounts of students' experiences with high-stakes tests (Anderson, 2004; Flores & Clark, 2003; Magnuson, 2003; Menken, 2006; Nichols & Berliner, 2007).

117

The following section reveals how English language learners experience high-stakes, diploma testing – in their own words.

Experience with the BSTs

Testing procedure

Unlike the TEAE, the BSTs are untimed, and students may take as long as they need to complete the test. Motivated by the high personal stakes of these tests, participants took advantage of the untimed nature of the tests, sometimes spending nearly the entire school day on one test:

> Arkadi: Yeah, I spend a lot of time with my reading test, it was like the third time and I wanted to pass it. So I was exactly in this room, with like, 8 people,
>
> K: Oh, you were in this room right here?
>
> Arkadi: Yeah, finishing… took me like 5-6 hours, I don't remember. (IG1_P5)

Arkadi had explained that the testing began in the gym, and as the overall group got smaller they were moved to a different environment – in his case an administrative conference room. There were no breaks allowed for lunch, and restroom breaks required an escort through the hallway. Stretch breaks are allowed, though in my observations of the reading test (administered to 78 students, ELL and native English speaking, grades 10-12), no break was given during the first 2.75 hours.[15] The test was given in one of the

---

[15] I was able to observe 2.75 hours while volunteering. After this amount of time had lapsed, 46 out of 78 students had finished the reading test, based on my participant-observer role during the administration of the BST reading retest.

school's gymnasiums first thing in the morning. Students were assigned four to a table, with little room to spare. The room was cold – many students had hoods on and arms tucked inside their jackets. It was also somewhat noisy: Overhead fans made it difficult to hear verbal instructions, even with the use of a microphone (Field Notes). Though all students took the BSTs in the gym, only Amina offered a negative assessment of the testing environment. She refers to the close quarters and her competitive nature, as she had in discussing the TEAE:

> Amina: I was competitive that time too.
>
> K: You were competitive, were you noticing other people finishing?
>
> Amina: Yeah, I was like oh, she's ahead of me,
>
> K: What would help that, Amina?
>
> Amina: Classrooms, like,
>
> K: Do you feel you were sitting too close to somebody else, when you took the test?
>
> Amina: It was like, right there, so like, I could see exactly what she's writing.
>
> (IG1_P3)

While the gym was used as a setting that could accommodate all testers, only the front half of the total space was actually utilized, with the minimum number of tables set up to fit 78 bodies. Overall, the human side of the testing experience and the impact of this test on students appears to have been lost, as the event was reduced to another task to be completed by already stretched guidance counselors, who appeared to be "running a

119

maze of technical procedures" (Macedo, 2006, p. xv). Most strikingly was the absence of a positive greeting or encouragement – nothing was done to rally the spirits of these students.[16] Instead, students were herded to their seats, given the instructions, and told to begin (Field Notes). Once most students had settled into the test, proctors (myself included) had to interrupt them individually in order to record the number of their test booklet (located on the front cover) (Field Notes).

Maria had taken more BSTs than any other participants, yet was particularly forgiving in her description of the testing procedure:

Maria: Well, I think the school's doing pretty good about it, cause uh, they give you all the time that you need, also, they give you like breaks,

K: They do give breaks, ok.

Maria: Yeah, so in 5 minutes you know you can stand up,

K: Is it enough, do you think?

Maria: I don't think it's enough, but I don't think they have to do that, even they do that, to help us, (IG3_P4)

Maria considers it a favor that the school offers her a break. In fact, the BST directions clearly states "breaks will be given every 30 minutes".[17] The policy for breaks was not

---

[16] Some students were new to the state and taking the test for the first time, while many others were repeat testers. All study participants were repeat testers.

[17] 2007-2008 BST Test Monitor Directions for Retakes: "You will be given as much time as you need to complete the test. We will take breaks every 30 minutes", downloaded from

followed for the reading tests, nor was it applied during the administration of the accommodated math test (Field Notes) during which native language accommodations were available.[18] The majority of ELL retaking this test was in the school's language laboratory using either a translated or English audio accommodation.

Like my observations of the reading test, observations made of the math test, while in a smaller setting, demonstrate the lack of ownership for this test by school staff, and the lack of consideration of and fundamental respect for test takers. In general, the retesting for BSTs seemed low in priority for the school. There was not enough staff assigned to proctor the math test, resulting in the evolution of my role as a participant observer to that of test proctor.[19] A substitute teacher assigned to the language lab later in the morning spent his time surfing the internet in the adjoining supply room.

The coordination of the technical aspects of the accommodation took some time (seat assignments by language group, corresponding test booklets, calculators) and

---

http://www.education.state.mn.us/MDE/Accountability_Programs/Assessment_and_Testing/Assessments/BST/BST_Manuals_Directions/index.html , 2-23-08

[18] The state of MN allows translated versions of the math test for ELL, but only provides translations for the most commonly represented languages (Hmong, Somali, Spanish, Vietnamese). Audio translated versions of the test are available, and in some languages, written materials are translated as well. Districts may have materials translated in other languages, but at their own expense. This often precludes many students from the accommodation. Students may request an audio CD of the math test in English.

[19] Because of my previous teaching experience in the school and with proctoring the BSTs, I was trusted and capable, though was never officially asked to do so.

required additional instructions for the use of technology.[20] These instructions were delivered by a teacher who was normally assigned to supervise the lab at that time of the day. The instructions were rushed and included no modeling or demonstrating – rendering parts of them unclear, even to me. Just as all students were finally settled and staff was ready to commence with the test instructions, the school-wide announcements began over the loudspeaker. One staff member immediately went to the office to ask that announcements be stopped. An assistant principal instructed the students reading the announcements to "speed it up". Finally, a district level administrator who was also helping with the testing went to the office and stopped the announcements. In the interim, at least two phone calls were made to the building secretary, who had apparently relayed the request to the head principal (Field Notes).

Accommodations

Participants in the study did not observe the school level factors surrounding the tests' administration. If they did, they did not share those experiences. They did have thoughts on the accommodations, though. Reactions among participants in this study varied in terms of experiences with and opinions of native language or English audio accommodations:

---

[20] In the language laboratory, translated audio tracks were loaded into a centrally controlled computer system, and assigned to students by their location in the lab. Once the test had begun, the students could individually control their audio, but the controls were visually different from typical audio CD players.

Li: Yeah, I did. And then, also, one thing is about, I think the math, they give it on the CD, which have their some people's like own language, in there, which, I don't think is fair, umm, fairs, because like, umm, there's one kind of questions in math, I remember, and it asking, what's the shape of the stop sign? and then, there's the English test, and then it's like A,B,C,D, like, kind of like _____ there, and then for me, it's like, I can't remember which word is it, because they all look similar, and then, for those people who have CD like, in their language, so they, can just pick up the answer right away, so it's like ...

K: So that was more of a vocabulary question, and if you heard it in your language you would know,

Amina: More even than in English

K: but as an English question it was vocabulary, not necessarily a math skill.

Amina: But we could've used the CD, couldn't we? They said that they had the English, the CD in your language, right?

K: In some languages…Was there a CD for Chinese?

Li: uhuh. (No) They had a CD for Somalia, and Spanish, and I think that's it.

(IG1_P3)

Li describes how, as a relative newcomer to English, a translated version of the math tests could have been to her advantage. Chinese is not a common first language among ELL in Minnesota. Neither is French – Genevieve's fist language:

Genevieve: I don't think it's fair, because I mean, I understand that everyone, come from a different background we all have different languages that we speak but when it comes to taking a test I think we should all take it in one language. Because having another person, who got a test, that is translated in their language, it is much easier for them to understand the question than the person who does speak another language, but has to take it in English because there is no way they could translate that for them. Cause basically it's kind of like having a person who understands totally what they're saying, what the test is asking to do because she hears it in her language, and you sitting there, you can't have it in your language so you basically have to figure it out in English.

K: So it adds another level.

Genevieve: Yeah, so it basically makes it harder for you because the stuff they might ask on the test you might know it, but you just don't understand the language, so basically I don't think it's fair, if they have to have translation, they might have to provide that for everyone, or if not, then just let everyone figure it out that way. (IG2_P7)

Both Li and Genevieve - whose first languages were not among those represented in translation – agreed that having the accommodation in some languages but not others is not fair. In the same group, on a different day, Winston (Liberian English speaker) agreed with the policy in principle. He felt allowing translations in other languages was OK, despite the fact he would not benefit directly from this. Jeanice disagreed:

124

Winston: I think it's fair.

K: You think it's fair?

Amina: I think it is, cause math has more like difficult words to understand.

K: Ok. Do you agree?

Jeanice: I don't know. But, I don't think it's a good idea.

K: Ok, why not?

Jeanice: Because, we are trying to learn English and improve English so, we just have to read it, in English and try to understand English so,

K: Even for the math?

Jeanice: Yeah. (IG2_P3)

Jeanice's sense of the accommodation policy being unfair differs from other participants' sense of the same policy. Li and Genevieve broach the issue of fairness from an equity standpoint: if the accommodation exists for some students, it should be available for all. Jeanice's view borders on assimilationist—if English is the goal, the tests should be in English. Having an English audio component available was generally viewed as not only fair, but helpful:

Sena: Cause on the CD they kind of explain it really slow,

K: Ok,

Sena: Yeah, they kind of take their time to explain it, for you, and kind of understand it more than when you like, reading, like, I don't know what to do,

K: Ok, so hearing it helps you understand, what to do.

125

Sena: Yeah.

…

Arkadi: It's the same for me like, hearing is much better for me than the reading and trying to translate for myself….Yeah, especially if it's a word which I don't know, maybe I hear it but I don't really know how, what it means, and when I hear it, I probably know it. (IG2_P5)

Genevieve, who disagreed with translated accommodations unless available to all, did feel that an English CD might help some students, even though it didn't help her:

Genevieve: Yeah I think that's a good idea because my, some people have a better understanding listening to someone than reading, so …(IG2_P7)

The English CD wasn't much help to Amina, though neither was the Somali version:

Amina: So I took the English one, and it was kind of going really slow, so I just went ahead of it, and then, in Somali I really don't know much of the words in the writing, how they, you know, the, (?) so I was like just took it off and just used my brain. (IG1_P3)

Amina describes her experience of using both a Somali and an English audio accommodation available to her, but how in the end neither proved effective. She had not learned mathematical terms in Somali, nor was she able to read the Somali script very well. The English CD was not really of use for her either. By her admission, it was too "slow". Even if students did not find the English CD helpful, there was general agreement that it was OK, and could help some students.

126

Lost Instruction

Whether testing in the language laboratory or in the gym, students participating in Basic Skills Testing were a small subset of the total school population – the rest of which proceeded with class as usual, as was the practice for the TEAE. Because of the untimed nature of these tests, far more instructional time is lost to testing than for the TEAE. Additionally, many students are retaking more than one of the BSTs, resulting in additional hours of lost instruction with each subsequent retest. Plus, seniors get one additional opportunity to retake each test during the spring of their senior year, resulting in further time away from the classroom.

> Edgar: Other people had classes, it was like a regular day but, we were just, um, testing and stuff,
>
> K:…so some people were testing, other people were in class. Do you feel like you missed out on anything by taking the test and not being in class?
>
> Edgar: I was happy I was out.

Academic disposition plays a role with these tests, as it did with the TEAE. Edgar admittedly appreciated an opportunity to miss class, whereas Arkadi and Belinda described the disadvantage this created in terms of missed instruction. For Edgar, the stakes were slightly different because of his status as a special education student.[21]

> Arkadi: I wasn't, I had two, (?) to make up

---

[21] Special education students are allowed additional accommodations if written into their individual education plans (IEP).

K: …Was it hard to catch up?

Arkadi: Yeah, especially with my English one we read the book, so, I was kind of a week behind. ..And I was like, totally out of it, and I thought, what are they talking about?

…

K: Ok. How about you, Belinda? Did you miss out on some class stuff?

Belinda: Yeah.

K: Did you feel like you got behind a little bit when that happened?

Belinda: Yeah, a lot.

K: Were you able to catch up?

Belinda: Yeah, after like two days. (IG1_P5)

Xoua's thoughts on missing class are tempered by her immediate need to pass the test. She is grateful for the opportunities for retakes.[22]

Xoua: Yeah, I missed like, 4 classes when I took the test, and I get behind but that's ok I think. I think it's a good thing that they let senior retake it. (IG3_P7)

There is some irony in the fact that students who are learning English and therefore need to make significant gains towards proficiency each year lose out on more instruction because of testing than any other students – in addition to days lost to the TEAE, they are more likely to be retaking BSTs. Retakes for these tests result in significant portion of a

---

[22] Seniors are one given additional retake opportunity for each test each year, beyond the two chances afforded all students. One of those opportunities for retesting occurs during the summer.

128

student's instructional day lost to testing. One could argue that the emotional and cognitive drain of taking a test for 4-5 hours results in the entire day lost in terms of instruction on each of the days a student is retesting. If the need for multiple retakes is a result of an inappropriate or poorly accommodated assessment for language learners, the tests begin to "undermine, rather than promote, the success of ELL" (TESOL, 2006, ¶ 4).

Perceived Rigor

Limited English proficiency is surely a factor in not passing, though academic readiness and disposition play a role as well, factors that often materialized in participants' expressions of perceived rigor. Expressions of perceived rigor on the BSTs varied among participants and appear to be influenced by a variety of individual characteristics, primarily success in passing the tests and number of resulting retakes for failing to do so. Several participants felt the tests were not hard:

Li: I don't think it was so hard, I just passed it first time, umm... (IG1_P3)

Genevieve: …I would agree that the BST might be pointless because basically kind of like, you know, all of the stuff we've learned probably elementary school and all that stuff,

K: So the level is too low?

Genevieve: Yeah. …and most of that stuff, people probably don't really remember them as well. (IG2_P7)

Li (a senior), and Genevieve (11th grader) were enrolled in Advanced Placement courses – an indicator of their strong academic abilities. Neither felt the tests were hard, and

129

Genevieve believed them to be at an elementary level. Belinda did not find them

challenging either, but lamented the length of the test:

> Belinda: It's easy but, like, it's long so you get bored.
>
> K: Ok, were they all long?
>
> Belinda: And if you know the answer, by the time you get to like to the 4th or 5th
>
> page you just guess because you get so bored.
>
> K: Ok, so you felt bored because it was so long?
>
> Belinda: Yes. And like I know if I read it I can get the correct answer but I get
>
> bored because so many answers and we sit there for like hours.
>
> K: And so it was kind of tiring? Was that, did you feel that way for the reading,
>
> the math and the writing?
>
> Belinda: Especially reading.
>
> K: Especially reading. Ok, so you started to guess after a while? Ok. Um, what
>
> was your experience with that test?
>
> Arkadi: Um, well, like, for example she said she gets tired, at the end, but for me
>
> it's the opposite, I will just tell myself to keep going, try harder, and going back
>
> and forth, and that's why I keep passing. (IG1_P5)

Neither Belinda nor Arkadi were seniors. Arkadi had passed all 3 tests, while Belinda

was still awaiting her scores on the math and writing tests. As a junior, the pressure for

her was not the same, as indicated by her admission of guessing out of boredom. A sense

of general academic disposition comes through their respective responses. Arkadi exudes

motivation and confidence:

> Arkadi: I thought I was prepared, and like, I was hoping that I would pass, I was
>
> actually believe that I passed it, so,
>
> K: Ok, you felt pretty confident?
>
> Arkadi: Yeah.
>
> K: Was the prompt something you were able to understand and to write about, the
>
> question they asked you?
>
> Arkadi: Yeah. I had no problem with that.
>
> K: How about you Belinda, for the BST this year, how did you feel?
>
> …
>
> Belinda: The writing I felt good, I don't know, I feel like I wrote a good story, and
>
> I used a good paragraph and stuff. So I'm confident but you never know.
>
> (IG1_P5)

Belinda is confident in her ability to pass, but with less certainty than Arkadi. Amina's

reaction (shared below) reveals some misunderstanding of the tests' content. In fact, the

reading and math are intended to be at an 8th grade level, not tests of discrete content

from 8th grade. Her response is somewhat contradictory, in that she subsequently admits

the tests are hard:

> Amina: I don't know why we have to take it, to graduate, because we are taking
>
> classes already to see if we improving, it's just to like, they just want to see, they

131

want you to, those tests, are from 8th grade, then you're a senior, you're not going

to remember all that stuff from 8th grade, so why bother taking the test, if you're

going to graduate?

K: Ok. So are you saying that some of the tests are easy?

Amina: It's kind of hard.

K: They're kind of hard. Ok.

Amina: I had to take a class for that. (IG1_P3)

The perceived level of rigor was high for Xoua, which in turn became motivating for her:

Xoua: Well, some of tests are ok, but the BST is really hard, and stressful. I don't

think it's boring at all.

K: Ok.

Xoua: Because we gotta think and, it made you want to study, for it, (IG3_P7)

In Xoua's case, a chance at college rendered the test motivating, though not without an

emotional toll.

Impact of Multiple Retakes

Pressure – and the resulting stress – increased by the number of retakes, and was

greatest for seniors not eligible for the exemption. Xoua – a senior – was in jeopardy of

not graduating because of the Basic Skill Test of writing. She was in good standing in

terms of credits and grades, and had even been accepted to a large university, contingent

upon her passage of the BST. She had attempted the writing test five times, and the

impact of multiple retakes is evidenced by her self doubt:

132

K: …how did you feel the last time you were taking tests?

Xoua: For me, the last time I retake the BST writing test, I was very nervous.

(IG1_P7)

In a later group Xoua shared:

Xoua:…I have a feeling I'm not going to pass,

…

K: I know. Oh you'll pass, let's think positive thoughts.

Xoua: I can't. (IG3_P7)

After attempting the writing test over and over again, Xoua found it difficult to be positive. She describes feeling at a loss for options after high school if she doesn't pass:

Xoua: Yeah, I don't know, like, when we start taking the test, if we don't pass it we still got next year, to this point, we just gotta feel like, if we don't pass, what do we do, you know? Because, you cannot go to college, don't you, just, gotta wait a couple of years, to pass the test, to get in or something? (IG3_P7)

Maria (also a senior) was feeling uncertain about her post-high school options, too. She wanted to go to cosmetology school, but needed a diploma to enroll. As the most seasoned re-tester of all participants, she believed more flexible paths to a diploma were warranted:

Maria: Yeah, I think some people, who cannot really pass it they can have another way to get diploma, or do the test. (IG3_P4)

When asked if she began to second-guess herself after awhile, Maria admitted that she did, and that after awhile, her nervousness would begin to impede her chances at success. She also spent a great deal of time on each re-test, claiming to have spent 4-5 hours on her latest reading test, and 4 hours on the writing test. Despite her repeated failed attempts at each of these tests, Maria had not given up and diligently approached each test with new determination, even though the effort was mentally and physically taxing. After working hard at the last writing test, Maria added; "I get headache" (IG3_P4). Remarkably, each of these students defies evidence that suggest such high-stakes can decrease motivation and lead to dropping out (Amrein & Berliner, 2003), or create disenchantment with learning (Wheelock, et al., 2000). In fact, both Maria and Xoua had dreams of continuing their education post-high school.

  Language acquisition a factor. Xoua and Maria independently identified language as an issue. Menken (2006) illuminates the challenges faced by students such as these, and describes how language places "ELLs at a serious disadvantage when test results are used as the primary criteria for high-stakes decisions such as high school graduation" (p. 522). This belief is reiterated by several other researchers and organizations, as well (Batt, Kim & Sunderman, 2005; Nichols & Berliner, 2007; TESOL, 2005). When I asked Maria if she felt there was a difference between kids who are not English language learners and kids who are English Language learners when it comes to tests, she agreed there was and had the following rationale:

Maria: Well the difference is people who have the second language sometimes they not really know which the word is, like um, like, if you reading the story but _____ some words, that come, you not, they have, like, a student who has second language sometimes they not really know. Which is the meaning.

K: Umhm.

Maria: So in that, that help a lot, from the way people, people who are first (?), like they really know, …So that, for that reason that white people pass it.

(IG2_P7)

In Maria's estimation, "white people" did not face the same barriers with language and could therefore pass the tests. This same phenomenon has been documented by the Center on Education Policy, which reports that white students are much more likely to pass high school exit exams on their first try than any other subgroup (as cited in Nichols & Berliner, 2007). The result for English language learners on a national level (many of whom are not white) is that, "we probably underestimate the abilities of millions of our youth with English-only high stakes tests" (Nichols & Berliner, 2007, p. 71). In a previous group, Maria had described her frustration with not understanding the writing prompt, and how this diminished her chances of passing before she had even begun.

Maria: It was stupid, I didn't get it,

K: You didn't get the question either?

Maria: And I asked for Ms K, and Ms. K came over and no, no I can't! Oh, Ms K I'm not going to pass it again! (IG1_P7)

135

As for reading, Maria claimed that if the test were offered in Spanish, she could pass:

> K: Ok, so in Spanish you can, you could pass that test, it's just the English version and the vocabulary.
>
> Maria: Yeah cause just sometimes I don't really get it, what the like, what's that, oh, what's that say? They confuse me. (IG3_P4)

Maria believed that allowing a translated version of the reading test would be fair. Minnesota requires students to pass a test of reading in English if they have been in the US three years or more – a not-so-subtle de facto language policy. Maria can read, but not well enough in her second language to earn a high school diploma. In Xoua's case, the differences between Hmong and English transferred to her self-acknowledged struggles with verbs in English writing, "…because when we speak in our language we don't really have like past tense, and future tense, and stuff like that" (IG1_P7). Because of the policy for exemption based on time, it is quite possible that English language learners with weaker overall academic skills than Maria's could earn a diploma because of their newcomer status, regardless of their overall language proficiency.

Because of their second language status (long-term ELL), both Xoua and Maria recognize their need for additional instruction in writing, and actually suggest a special course for students in their situation:

> Xoua: Yeah, they have class like, during the summer, but why don't they have class during like, the school time, like this, when like student could just like, take the writing class as like, a class, and get a grade for it,

136

K: Would that be something you would be interested in as well? You mentioned writing was hard for you, do you wish there was more of a writing focused class?

Maria: Yeah, even a, I went to summer school, to pass, and I didn't pass it.

(IG1_P7)

Unlike many other schools in Minnesota, Aldrich High School did not offer courses specific to remediation for BSTs during the school year, but these students are actually asking for one! In the absence of such a course during the school year, these students were left to face the high stakes exams with little systemic support. Though because they were seniors, an arrangement was made for several hours of tutoring before the final scheduled retake; an opportunity in which Xoua participated. Because these tests are not tied to systems accountability, the same degree of urgency does not exist as in places where schools have become focused on test prep (Bernstein, 2004; Menken, 2006)

At the time of data collection, Xoua and Maria were anxiously awaiting their latest retake scores:

K: …I bet you are nervous about the BST.[23]

Xoua: Yeah, I am. I keep asking Ms. ? (unclear), is it in yet? She like, no, but it will be, soon,

In Maria's case, impending graduation weighed heavily on her response:

K: How's that make you feel while you're waiting?

---

[23] At the time of data collection (May), seniors were awaiting the results of their latest retakes.

Maria: Hm, that make me feel stupid because I didn't even know if I was going to go, I haven't, to be honest, I haven't, buy anything for the ceremony thing, whatever, Ms. ? told me um, I can go to the ceremony because I have, all my credits and everything, but why I'm going to go,

K: So you don't want to even go to the ceremony unless you pass the test?

Maria: Yeah.

K: Ok. So, do you think, will there still be a chance to buy the gown? If you find out you've passed?

Maria: I hope so, I don't know. I don't really care about the ceremony, I just want to just, get my diploma and that's that.

K: Ok, are you feeling frustrated,

Maria: Yeah, I'm really, (extended pause)

K: Tired of the whole process?

Maria: Yeah, I'm really tired. (IG3_P4)

For both Xoua and Maria, the emotional impact of not yet having passed the BST was high. They had not given up completely, but were clearly weary of the retakes and affected deeply by the consequences. After a short discussion, I clarified with Maria that she had, in fact, taken the BST reading test nine times, and was close to passing on more than one occasion:

Maria: yeah, cause, cause, some scores I was really close to passing, and yet I don't pass it I don't know why.

K: Yeah, you had a 596 in 2005, 4 points away. (IG3_P4)

Maria was visibly and audibly dejected by the experience of taking the reading test so many times. Since first taking the test in 8th grade, she persistently participated in each subsequent opportunity to retest afforded to her, including three consecutive summers. At the time of the data collection for this study, Maria had only passed the math test (in 2005). Her previous eight reading scores ranged from a low of 547 to 596 (600 is needed to pass). She had taken 18 Basic Skills Tests in total since 8th grade. As minimum competency exams, the policy behind the BSTs is not for accountability to "leave no child behind". This lack of system accountability is at the expense of students like Maria and Xoua, and their prospects for furthering their education beyond high school.

As seen from the data presented above, the knowledge participants had of tests (research question one) and the experiences they described (research question two) vary considerably based on the perceived stakes. The process of coming to knowledge was often based on individual and shared experiences, and conversely, experience was often influenced by individual and shared knowledge. Language proficiency was a primary factor in either facilitating successful passage of the diploma tests, or in creating an obstacle which ultimately led to increased frustration, anxiety, and multiple retakes. These data support what other researchers have determined: language proficiency seriously disadvantages students on high-stakes diploma tests such as the BSTs (Anderson, 2004; LaCelle-Peterson, 2000; Magnuson, 2003; Menken, 2006; Nichols & Berliner, 2007; Quest, et al., 1996).

139

I now turn to a deeper exploration of the role of students in the testing process based on the overall sense they make of statewide testing (research question three).

Making Sense of Testing

TEAE test is to walking as the BST test is to running.
(IG2_P3 / AT2_Amina)

Knowledge and experience are key components of the sense-making framework borrowed from Spillane, et al. (2002), which is why they were useful for organizing the previous analysis, but they are not the only components. As demonstrated earlier in this chapter, knowledge and experience transcend individual cognition. The data related to the first two research questions illuminate how multiple factors influence the sense students make of testing. These factors ranged from individual cognitive and emotional characteristics that affect the ways in which students view testing, to the influence of the research process itself in constructing new, negotiated understandings and characterizations of experiences. A further exploration of the multi-dimensional nature of sense making is presented here, without isolating data pertaining to knowledge and experience from that process. Rather, I return to the framework mentioned previously and loosely explore the processes of sense making from a three-tiered perspective: those instances that are individually situated (tier 1), socially situated (tier 2), or influenced by the policy stimuli (tier 3). I take a far more integrated approach to the framework,

140

however. My data reveal intersections of factors that preclude exclusive assignment to one tier or another.

Similar to findings related to students' knowledge of and experiences with statewide accountability tests, the sense they make of the tests varies by the tests, too. Therefore, findings that demonstrate how English language learners make sense of testing (research question three), will be presented first by each of the tests (TEAE, BST), concluding with an attention to testing in general.

Making Sense of the TEAE Test

The examples of knowledge shared and created around the TEAE presented at the beginning of this chapter underscore the complexities inherent in the current policy for this test. The process of coming to an understanding of this test was amplified by individual factors as well as peers, teachers, and in the case of this study - the researcher. Additionally, there were policy signals - the filtering of the TEAE policy's intent through teachers along with signals sent by the design of the test. Combined, these factors become moderating variables that help to explain why students' sense of testing varies (Louis et al., 2005).

Individual factors

Individually, factors such as time in country and overall academic disposition played a key role, as did the stigma for some students of being required to take the TEAE. For example, Arkadi's disposition towards school and anything related to it is very no-nonsense and academically oriented: "TEAE test was for me like any other test. I

always take any test very seriously and try my best" (IG1_P5). Arkadi had accepted

testing as part of his school experience, and did not question it as an institutional practice

(Koyama, 2004). Time in the US played a role, as well. Those participants who were

relatively new to the US had a greater tolerance for the TEAE, while those whose English

proficiency was stronger and/or had been in the US longer had far less patience for a test

they deemed "stupid"(IB1_P7_Maria), or a "waste of time"(IG1_P3_Amina). Similar

adjectives were used by participants who viewed the TEAE as being too easy when

describing how the test made them feel. Xoua claimed the test made her "feel stupid,

because that seems like first grade" (IG1_P7).

An unintended consequence for many proficient (or nearly proficient) ELL was

the stigma associated with taking a test of language proficiency; the very act of requiring

students to take a test of language proficiency implied doubts about that proficiency in

the minds of students. Regardless of whether such doubts about a student's proficiency

are warranted, if the student is told she no longer requires ELL classes, that message is

contradicted come testing time, by peers and teachers alike. In this sense, the 'flexibility'

built into NCLB is very clearly designed with the system in mind. Allowing more

proficient students to test (in order to demonstrate growth) serves the interest of school

districts, and reinforces or "reifies" the position of these students as language learners.

While it may be "theoretically possible for students to change positions, this is, in fact,

almost impossible to achieve, given the border created by the policy itself and the policy

appropriation process" (Koyama, 2004, p. 403). As such, the "assessment policy and

practice is complicit in defining ELL students as deficient and uni-dimensional" and

becomes a framework that for these students leaves "little room for celebration but rather

a persistent focus on identifying a particular limitation (often read as a deficiency) to be

remedied" (LaCell-Peterson, 2000, p. 35).

Social factors

There was a general aura of uncertainty around many aspects of the TEAE. Much

of the sense participants had of the TEAE prior to discussions in the inquiry groups was

admittedly influenced by their peers, as well as teachers. Messages from teachers, while

sometimes contradictory, were both informative and influential. Jeanice shared how staff

informed her of the need to take the TEAE test after she had been in the school one week:

"I said wow, ok! I will just go in, I can say for me I was really slow" (IG1_P3).

Winston explained how a particular staff member instructed him to take the test:

> Winston: Cause um, they told me, Mrs. K told me.
>
> K: Mrs K told you, ok, so, did you get a pass then too, or did they make
>
> an announcement?
>
> Winston: I had a pass.
>
> K: You had a pass? Ok. Did anybody explain why you had to take it, even
>
> though you're not in an ELL class?
>
> Winston: No. I just took it. (IG2_P3)

Participants had a profound acquiescence for anything they were instructed to do by Mrs. K.,[24] as evidenced by Winston's and Jeanice's examples. The authoritative nature of staff affiliated with the ELL department, as described by participants, led to an uncritical acceptance of the need to take the test and is indicative of an "ingrained passivity" in students (Corbett & Wilson, 1995), even if new to the US. The exception to this was Genevieve's example of contradictory messages; one staff member was, in her word, "running" after her, while another told her she didn't have to take the test (IG1_P7_G). Participation in the test cannot be equated with *engaged or willing* participation, however.

Policy Signals

In addition to the policy for 'flexibility' described previously (allowing districts to test re-designated students for 2 years), other aspects of the TEAE itself sent clear signals to participants. Lack of complete engagement with the TEAE was in part due to the design of the TEAE. Regardless of proficiency level, the test begins with pictures and basic vocabulary, eventually building complexity with each subsequent section. This design sent clear signals to participants, to which they had not only strong reactions (described under research question two, above) but well articulated suggestions, as well.

---

[24] Mrs. K. is an educational assistant in the ELL program at Aldrich High School.

144

In response to a written prompt asking what people who make tests should do, participants in one inquiry group had mixed reactions. Winston (Liberian English speaker) felt the test should be more challenging, whereas Jeanice had a different reaction:

> Jeanice: Hmm, for me, I say that they should just give the test, like that way, because we, we're different, we can say that it's not challenging, and another person will say no it was hard, so, they should just keep it like that …Yeah, because they like, student like, when I just came, last year, I just did like one week in school and then I had to take the test, so, that is hard when you don't speak English, and you just came like that, and so, that's why you should just stay like that ...
>
> K: Ok, so not be too much more challenging for students in your situation who are just new to English,
>
> Jeanice: Yeah. (IG2_P3)

After further discussion, the group came to a consensus that perhaps students should be tested according to their proficiency level:

> Arkadi: They should just make tests for people ... For certain groups.
>
> K: Make them for certain groups?
>
> Jeanice: They should make a test like, for a student maybe who have 5 years, or 3 years, or depending on the years you came here.
>
> K: Ok, so depending on years,

145

Arkadi: The level you're in. **(IG2_P3)**

Arkadi offered the same suggestion in a different group, "put the test at the same level as the kids are" (IG2_P5), as did Genevieve:

Genevieve: People who make tests? Hmm, they should make it harder, or easier depending on the student level of knowledge.

K; Ok, so if a, let's think of a language learner for example, somebody who's new to the country doesn't speak a lot of English. They should get an easier test, in English?

Genevieve: Yeah, well, not an easier test, but they should get a test that is, aportable to their level of English. To see actually how they are doing. You know?

K: And then if you're in a higher level, or higher grade,

Genevieve: You should get something more challenging because you know more than the other person, so, (IG2_P7)

When asked if she felt that is how the test is constructed now, Genevieve referred to a comment made by another group participant, "it's not really, for instance, the TEAE test, like, um, what's her name, she said they had questions in there like, kind of like, kindergarten questions, …So it's basically ok, at this level you should give me something that will make me think not something that will just like, could do it with my eyes closed, you know. (IG2_P7). Genevieve makes her opinion clear that tests should be at the right proficiency level, as well as academic level. In other words, a "one-size-fits-all" approach

146

to testing does not work for an ELL test any more than it does for tests given to all students (Menken, 2006; Nichols & Berliner, 2007; TESOL, 2005). In fact, the current iteration of the TEAE does not recognize the heterogeneity of the ELL subgroup, and neither does NCLB (Batt et al., 2005).

As these participants have demonstrated, the diversity of experience, knowledge, and individual characteristics within the subgroup of students labeled ELL is vast, and worthy of consideration in the crafting of testing policy and the resulting tests. Participants recognized and described how testing all ELL (and recently re-designated ELL) with the current iteration of the TEAE had some flaws based on the design of the test. Additionally, participants' general lack of understanding of the purpose of the test raises questions of engagement and buy-in, as well as the degree to which teachers, administrators, and others recognize students' agency in the testing process.

After having time to reflect on and compare the TEAE to other tests, participants were able to develop a clearer understanding of the policy for this test, as well as propose thoughtful, relevant changes in the policy. They began to develop, though minimally, a sense of agency, belonging, and competence (Mitra, 2004) as participants in these inquiry groups. My own role in the process of negotiating understanding of basic features of the TEAE helped create the framework within which they could do so. When discussion and questions turned toward the BST, my researcher role had less to do with explaining or clarifying, and more to do with listening and asking probing questions. Because the

147

stakes are so high for Basic Skills Tests, participants had a far more sophisticated level of reflection and interpretation of these tests that was also highly personalized.

Making sense of the BSTs

Personal experience with not passing one or more of the tests had a tremendous influence on framing individual responses. At the same time a collective spirit emerged; those who had recently passed all three tests[25] also had a shared sense of some basic inequities. These inequities were influenced by messages and signals sent by peers, teachers, and the test policy itself. The influence of these factors was multidimensional. In inquiry groups where the passing rate among participants was mixed, students who had passed were extremely respectful of their peers who had not, and waited until the end of an exchange to offer a contradictory response, often with great tact and diplomacy, couched in personal experience. This was seen in the discussion of students' experiences with both types of tests. For the TEAE, participants newer to the country (Jeanice, Li) had more positive experiences than others with the TEAE as a test of their language proficiency, and said as much (IG1_P3). On the BSTs, students who had successfully passed the tests (Li, Arkadi, Genevieve, Winston) did not flaunt their academic skills in discussions around the challenges of the BSTs for others in their respective groups.

Perceived inequities

---

[25] All participants had retaken at least one BST, and could relate to those who had not.

Conversations around the BST were often wrought with a tinge of unfairness. Perceived inequities emerged around several topics, namely systemic support, the test requirements, the arbitrary nature of scoring and the subjective nature of the test prompts.

In the following sample of dialogue, Xoua identifies several perceived inequities related to the degree of support offered students at Aldrich high school:

> Xoua: …I think the BST test is unfair, because I see a, like, some students they barely pass their classes, but they pass the BST, and they have like "A" [Alternative] school to help them graduate, even though they miss so many credit, and we have all the credit we need to graduate, but we only have like, for example me I need to pass on writing test, and, we don't really have classes to help us pass that test, and, I think it's not fair because um, some student do real well on testing and some don't, so, they should depend on like, school and testing too but they should not, just like, if you don't pass then you can't graduate because I think that'd be like, too much pressure on their like, second language class, because when we speak in our (language?) we don't really have like past tense, and future tense, and stuff like that. (IG1_P7)

First of all, Xoua points out that at Aldrich high school, students who are in jeopardy of not having enough credits can get help through a special program within the school – "A" (Alternative) School. Students like Xoua, who have fulfilled the credit requirement but have not passed one or more of the tests, do not have a similar level of support. It is as if the school has absolved itself of further responsibility for providing instruction in reading

149

and writing beyond that offered through the ELL department. Mitra summarizes how student voice research can force administrators to pay more attention: "Through open conversations about injustices in schools, student voice can raise equity issues that tend to get swept under the rug by administrators and other adults in the school who would rather avoid controversy" (2006, p. 653). Mitra also describes how involving students (especially those who are failing some aspect of school) forces schools to assess the problem in the system, rather than shift blame onto the student.

Xoua described having taken a summer writing class, where she earned an A, though her teacher pointed out the need to continue to work on verb tenses. Xoua recognizes her continued struggles with certain grammatical features of language, which she feels are influenced by her first language, Hmong, and was frustrated by the lack of a class that would specifically address her writing needs during the school year. If both Xoua and her teacher recognize her struggle with verb tense, where is the instructional follow up? Xoua was acutely aware that such a plan was lacking. She also points out that tests are one part of schooling; not everyone does well on them. In a later inquiry group, Xoua reiterated that some ELL do better than many native English speakers in some areas of school and that length of time in the US and other factors can be more influential than language in determining success on tests. While she admitted her language was still a struggle for her, she believed other factors were as much at the core of why she had not passed the writing test. Her comments illuminate "the research –policy schism", that until it is bridged, "LEP students' achievement test scores should not be used to deliver high-

stakes sanctions but rather to make diagnostic decisions about how to better support these students" (Batt, et al., 2005, p. 11).

The diploma tests were not tied to NCLB or adequate yearly progress at the time of this study. The fact that Xoua felt little systemic support for passing this test speaks to the unfair burden of high-stakes tests on students learning English, accentuated by the overall lack of ownership of this test on the part of the school. Nichols and Berliner warn of a new type of discrimination creeping into schools against "score suppressors" – those students whose scores end up damaging a school's reputation. While the BST scores had little overall effect on Aldrich High School's reputation at the time of this study, the relative laissez faire attitude towards the test despite the ultimate consequence to students hints at a more subtle form of the same phenomenon – one likely to become more overt when a diploma is tied to NCLB accountability tests (as is the case beginning with the 9th grade class in MN in 2007-2008).

Requirement of a diploma test. Several participants believed strongly that passing classes and earning the requisite number of credits should be enough of a requirement for a diploma. For Maria, this sentiment was influenced by her interaction with Mrs. K, an educational assistant in the ELL program:

> Maria: I was talking with Ms K and I told her that, why, if I have my grades, and in everything, like doing well, why, I'm not gonna get my diploma. And well, … Even Ms K said, I know it's kind of stupid, but… (IG1_P7)

151

Belinda expressed a similar sentiment; "Because, some people are smart they get good grades, and stuff, and if they fail, the, Basic Standards Test[26], let's say they pass the reading and the writing, and they fail math, they not going to graduate, and all their hard work, just, down to the toilet" (IG1_P5_Belinda). She articulates the possible reduction of four years of hard work to the consequence of failing just one test. Additionally, she was unable to move out of her 'LEP position' because of her math test, not because of her language proficiency (Koyama, 2004, p. 412).

Genevieve could not quite commit to an opinion of diploma tests. In her case, she was an 11th grader who had passed all three BSTs, and two other participants in her group were seniors with multiple retakes. The research setting seemed restrictive for Genevieve due to the makeup of the inquiry group, and her respect for her peers:

Genevieve: Well, I mean, the BST, I would have to agree with them, cause, basically, well, I'll just say, I don't really agree with them, but I do agree with them…(IG1_P7)

She went on to describe factors such as learning English, having enough credits, and the pressure associated with graduation tests as reasons why the policy is unfair. Her overall comments, however, indicate her view of the tests as more of an annoyance, than a barrier; after talking about meeting credits, taking care of college applications, scholarships, etc. she noted "but then you have to come back and think oh no, I have to

---

[26] Reference is sometimes made (by participants, and generally) to BSTs as Basic *standards* tests, remnant of a two-tiered system of standards and assessment in Minnesota: Basic and High Standards

do this, before I get out of here" (IG1_P7). In terms of her participation in Advanced Placement course, she asks:

> G: …I'm taking college courses, why would I need the graduation requirement?
>
> K: Ok, so it's a higher level than the basic test.
>
> Genevieve: Yeah. so basically it's kind of like, I'm challenging myself to take this class and I have a certain grade in this class, so why do I have to take this to graduate? (IG2_P7)

Overall, many participants were far more motivated by the need for credits, and their GPA and overall academic performance in school, than by the Basic Skills Tests. When asked whether colleges pay attention to the test scores, Jeanice offered, "Not the BST, I think they don't care about the BST" (IG2_P3). Other tests like the ACT or SAT were mentioned as being more relevant to post secondary institutions.

Seniors, especially, who had taken the tests multiple times questioned the overall fairness and efficacy of graduation tests if they were doing well in school. For Maria, who was fearful of not graduating, this was expressed clearly:

> Maria: Well, um, for me, my opinion is, why do we have to take those? Well, sometime I think it's not fair because um, since I am doing really well in school. You have everything, you have your grades, everything, even you don't pass one of them, but you don't get your diploma. So I'm like think, it's not really good for a lot of students. (IG1_P7)

Xoua approached the topic of fairness from a slightly different angle: the fact that not all states have the same requirement:

> Xoua: I have a question for you, like, you cannot go to the U of M if you don't pass the BST, but why do U of M accept student from Wisconsin? Where, when they don't need to take the BST too? … Cause, I know a lot of people from Wisconsin go to the U of M. I was just wondering, if I don't pass, I can't go, but they don't take it, but they get to go in, so… (IG3_P7)

In raising this question, Xoua points out the relatively arbitrary nature of graduation test requirements when considered within a national perspective, when roughly half of the states do not have a high school exit exam.

Arbitrary nature of test scores. The cutoff score for the tests was viewed as arbitrary as well. I clarified with Maria that she had, in fact, taken the BST reading test nine times, and was close to passing it on more than one occasion.

> Maria: yeah, cause, cause, some scores I was really close to passing, and jut I don't pass it I don't know why.
>
> K: Yeah, you had a 596 in 2005, 4 points away.
>
> Maria: Yeah. (IG3_P4)

Maria was visibly and audibly dejected by the experience of taking the reading test so many times. Her previous eight reading scores ranged from a low of 521 to 596 (600 is needed to pass).

154

Sena expressed similar frustration at the seemingly arbitrary score needed to pass, a frustration shared by her friends.

Sena: …My friends and I were talking about the BSTs just a minute ago. We kind of felt like you know some kids couple of my friends got like 599, and we don't think it's fair because then they have to take it all over again,

K: They missed by one point.

Sena: Just one point, so we were thinking, why can't they just keep their score, and then take the test again and add up your score?

…Cause I got 597.

…

K: That's frustrating isn't it?

Sena: It is, I cried. (IG2_P5)

Fortunately for Sena and her friends (all 11th graders), another year of opportunities for retesting awaited them. For those who fail by just a few points, "high-stakes testing makes education a cruel endeavor" (Nichols & Berliner, 2007, p. 71). Cruel or not, all participants, regardless of grade level, awaited the arrival of scores for the BSTs anxiously.

K:…were you waiting for those to come?

Amina: Oh yeah.

Jeanice: uh-huh. like, you like, oh my god, I can't open it, (IG1_P3)

155

Arbitrary or not, all participants described the anxiety they experienced while waiting for their scores. Even Jeanice, who qualified for the exemption and knew it, expressed anxiety around scores that was amplified by participants who had come close to passing one or more of the tests multiple times, but had not yet passed. A sense of arbitrariness also applied to the topic of the writing prompt from year to year.

> Xoua: Well, for the BST writing, I think students should write their own essay. What they want to like write about. Because it would be easier for the student, because some topics more like, if you, know about the topic than you'd be better at writing about it, and some student don't really get it, and so it'd be harder for them to write an essay about it. (IG3_P7)

Both Xoua and Maria expressed frustration with the latest writing prompt they had faced. They each put an enormous amount of effort and time writing an essay which they were uncertain was addressing the topic, adding to their already increasing anxiety. Participants in multiple groups suggested that when it comes to testing writing, students should be given a choice of topics – a practice which would, in fact, reflect good pedagogy. As Xoua noted (above); "If you know about the topic, then you'd be better at writing about it". Background knowledge, tied to opportunity to learn, impacts the success of students with testing. Uncertainty around the writing prompt added to the anxiety and sense of unfairness for seniors retaking the test.

Making sense of tests, generally

Interestingly, all participants described significant, high stakes exams in their own countries (where exams and curriculum are often national) and were relatively uncritical of the practice. They are not opposed to the idea of having to take tests, but based on their knowledge and experience with testing in the US, and the perceived inequities described above, they struggle with the idea of needing a test to graduate in the US if they are passing their classes. Participants also struggle with the idea of a language proficiency test that is undifferentiated, with no real perceived stakes for students.

There is a general sense among participants that if tests have a clear, relevant purpose, and if the tests themselves are fair and given at an appropriate level, they are acceptable and more likely to be taken seriously. Perceived stakes ultimately matter, as Li describes in her response to the BST: "because it affects you to graduate or not, so, they will take it more serious, and then, just, it's like better than TEAE test, you know, nobody cares" (IG1_P3). But as described previously, high stakes alone do not make a test fair, and in many cases, this raises questions about the validity of the resulting score. For one thing, the opportunity to cheat is perceived as real:

> Arkadi: I just know like, students, skipping one day, and then asking for answers the next day from different students,
>
> Belinda: Do they do that? That's bad. ..Um, I was gonna say, even when you take BST test, it's easy to copy too.
>
> Arkadi: Well I mean like, if you copy you never know if it's right or wrong.

157

Edgar: Yeah, but you gotta know who.

Belinda: I'm not saying we do copy I'm just saying it's possible. (IG1__P5)

Nichols and Berliner have, in fact, documented the extent to which the pressure to do well on high-stakes tests has created an educational system that encourages cheating (2007). To reiterate, participants do not believe tests are inherently bad. In response to the sentence starter "I think tests…" on a written task (AT_2), participants alluded to the value of formative assessments which can provide immediate feedback to inform instruction and learning:

Arkadi: Uh, I put just a way to show yourself, how well you learn the topic.

…

Sena: Um, I think they are very important to let teachers to know how you're doing, and your, like your weakness, like stuff like that. (IG2_P5)

Maria: Um, I think tests are not really good, cause um, well, depends what kind of test, if there's for math, or like kind of for regular classes, sometimes I think they're ok, that way the teacher they know how you go learning,

K: Ok. So do you mean like tests in class, are ok?

Maria: Yeah, but the like, BST and the uh, another test, oh, which test is it?

K: The TEAE test?

Maria: TEAE test? Nah, I don't think that really good. (IG3_P4)

The same value in terms of formative feedback was not as evident to participants for statewide assessments. For one thing, there had historically been a significant delay (6 or

158

more months) in receiving scores for the TEAE test, rendering them useless for providing

guidance to teachers (Chabran, 2003; Magnuson, 2003; Nichols & Berliner, 2007). In

terms of the BST, the only aspect of the score anyone was interested in was whether they

passed or not. Whether a student barely passes, or passes by a landslide was irrelevant to

students, school staff, or parents of these participants, many of whom were in the dark

about most aspects of the tests that had such a great effect on the students. It is incredibly

difficult to "justify that some point on a scale really separates out those who understand

something from those who do not" (Nichols & Berliner, 2007, p. 17). Not only are the

cutoff scores arbitrary, they are also intended to represent achievement at a *minimum*

level of education (Magnuson, 2003), a point that becomes lost when the focus is on

either passing or failing.

Students as Knowledge Brokers

Participants in this study described their role as knowledge brokers in terms of

their role in interpreting test scores to their parents. Arkadi described how it is his job to

explain test results to his parents:

Arkadi: Um, I try to explain to them, I mean they had trouble understanding some

of them, but they kind of understand that it's important, so, they know I'm doing

my best. (IG3_P5)

159

Jeanice, described her frustration responding to her dad's questions about what the scores mean: "And I said, I don't know too!"(IG3_P3). In a later group she indicated her parents' desire to understand:

>Jeanice: Umhm, yeah, like my parents they really want to know what's going on, if you don't work very hard they will be so, I don't know, it's kind of sad.
>(IG2_P3)

Jeanice's parents, like many parents of ELL, are often in a position to receive important information via their children; information their children are often unable to explain clearly. Amina gives the example of reducing tests to either important, or not: "We go through it. She just thinks that I have to do good, to graduate, and I just tell her that some tests I don't and some I do" (IG2_P3). If the test does not impact graduation, it becomes reduced to a further state of irrelevancy by the time its message has been interpreted through students. In Xoua's case, even the BSTs don't warrant a full explanation to her parents:

>Xoua: My parents? To tell you the truth, they have no clue what we've been, like what we talking now, like
>K: Ok
>Xoua: they don't even know what's the BST, we don't, we don't tell them about it, because they would not get it, you know they would just be like, ok, a test,
>(IG3_P7)

160

Only one student, who had an older brother go through the school, felt like her parents had a relatively good understanding of the tests, especially the BST, whereas Maria's siblings (with whom she lived) had passed BSTs. They not only understood the consequences, but added pressure as well:

> Maria: They say, well my brother's kind of supporting me, but since I, it's like, I can't believe you haven't passed it, like I can't, it's easy, just read it, and answer question that's it. why you don't pass it like uh!
>
> K: So it's hard for him because he passed it.
>
> Maria: Yeah he passed it. The first time he passed it.
>
> K: Did he?
>
> Maria: Yeah, my sister too, it's like, why you don't pass it? I don't know!
>
> (IG3_P4)

With parents, as with students learning in schools, prior knowledge (in this case with schooling in the US) and language proficiency are enormous factors when it comes to comprehending the results and impact of statewide assessments. In turn, the roles students lay in brokering information to their parents affects the sense they make of testing.

Testing Fatigue

There was a general sense of testing fatigue among many participants, which contributed to the overall sense many made of the tests. In the literature on students

perspectives and testing, this is addressed more in terms of test stress (Anderson, 2004; Nichols & Berliner, 2007; Wheelock et al., 2000) than in terms of sheer number. Participants in this study reacted to the amount and length of testing, and the sheer endurance required as a result. Jeanice indicates that statewide assessments aren't necessarily bad; "I think, umm, in one hand, it's a good initiative, but I think we don't have to take some of them too, but, I think it's not necessary, because it makes people really stressful and tired"(IG2_P3). Jeanice alludes to the number of tests required, while Belinda and Arkadi had comments on the length of some tests:

> Belinda: Don't make it long.
>
> …
>
> Arkadi: I kind of agree, because when you're working and you get tired, and you kind of don't care about it, you want to go to your next class and go home.
>
> (IG1_P5)

Arkadi, who was by far the least critical participant of his testing experience, expressed a frustration that can develop when fatigue sets in, along with a desire to "go to your next class". Missing classes, as described previously in this chapter, contributed to overall testing fatigue, reiterated here by Genevieve and Jeanice:

> Genevieve: Yeah, they still missed class. So basically, if we have to have tests like that, they should actually kind of like, approximate the time that it's gonna take, like almost every student, and then, provide that time so everyone would just get it over with, and not, miss their classes, …(IG2_P7)

162

Jeanice: I'd say that, maybe they should choose a day when people are not going to school so the other students are not miss the class, so yeah. (IG2_P3)

Many aspects of a test's implementation are left to the discretion of the school within systemic constraints. At this site, the TEAE and the BSTs were administered during the regular school day, without suspending classes (or even announcements). On occasion, participants in this study made reference to the Minnesota Comprehensive Assessments (MCAs)[27], the tests of academic achievement used to demonstrate adequate yearly progress under NCLB. For the MCAs, students were assigned to a classroom, a practice Amina suggested would be good for all tests:

Amina: Give it in a place where they feel comfortable in.

K: What would be an example of a place where you would feel comfortable?

Amina: Class. I mean, the MCA, I took it in a class I felt more comfortable there than in the new gym. (IG2_P3)

Although the MCAs are given to an entire grade level at one time, and the TEAE and BSTs to smaller numbers of students, there is no reason they, too, could not be administered in classrooms. Minimally, students should not have to bear the burden of lost instructional time. This practice will become more equitable as Minnesota transitions to a new system of diploma testing. Beginning with ninth graders in 2006-2007, students

---

[27] The reading assessment is first given in grade 10, and math in grade 11. A science test is given during the year in which a student takes biology. These were not focal tests for this study, because not everyone had taken each test.

will still have to pass tests for a diploma, but the requirements are changing. Students now have to pass a test of writing (first administered in 9th grade), as well as diploma requirements embedded in the 10th grade MCA of reading and the 11th grade MCA of math. The tests used for school accountability purposes under NCLB nave now become high stakes for students, and at a higher level of academic standards. After sharing this change with Genevieve, she had some advice:

> Genevieve: I don't think, I don't know, thinking that ELL student don't have the right, well, not that they don't have the right level of English to be able to take the test, but I'd say ELL student starting maybe from 2 or down, or level 2 to down, shouldn't be able to take that,
>
> K: Ok, so maybe by level,
>
> Genevieve: Yeah,
>
> K: If they're really struggling with English, they shouldn't have to take the test.
>
> Genevieve: Because that would not benefit the school at all.
>
> K: Right, because you know they're going to score low.
>
> Genevieve: Yeah. (IG2_P7)

Genevieve recognizes that scores on these tests will impact the school directly, yet also recognizes the overall futility of requiring an English language learner with low proficiency to take a test of academic content that was designed for native speakers (La-Celle Peterson, 2000; Magnuson, 2003; Menken, 2006; Nichols & Berliner, 2007; TESOL 2005). As Magnuson notes, "ironically, it is the fear of not treating each student

164

fairly which seems to lead us into not recognizing that different students have different needs" (2003, p. 213).

Finally, I asked participants whether they had ever been invited to discuss educational policies beyond the classroom level, and the consensus was a resounding 'no'. When asked if they would like to have a say in policy discussions in the future, especially regarding testing, one group responded:

Belinda: Of course.

Arkadi: Yeah, probably.

Belinda: I say yeah, yeah, yeah! (IG3_P5)

Participants in this study, as were students in other studies of school policy or reform (Mitra, 2004), were grateful and enthusiastic for the opportunity to be part of the conversation.

As the findings presented in this chapter demonstrate, students' knowledge of and experiences with statewide testing play an important role in the sense they make of tests. These English language learners are thoughtful, articulate, and eager to comprehend and comment on the policies that impact them. Li's quote at the beginning of this chapter is indicative of this. Students are key stakeholders from whom others have much to learn.

CHAPTER 5

CONCLUSION

We recognize that there is a long way to go before the education establishment regularly accepts adolescents' viewpoints as anything more than either the ramblings of immature youth or statements from the endearing but naïve.

(Yonezawa & Jones, 2007, p. 708)

Beyond contributing to the fields of educational policy and student voice initiatives, the intent of this research is to spur advocacy on behalf of and in conjunction with English language learners. This research documents in students' own words the sense they make of statewide accountability testing in Minnesota. Whether for systems accountability (the TEAE test of language proficiency), or for individual accountability as a requirement for a diploma (the BSTs), the findings reveal both the complexities of and urgency for exploring testing policies from students points of view. Additionally, findings validate the need for students' voices in policy research, and suggest several implications related to their experiences with statewide accountability testing which will be explored in this chapter. Limitations of the current study will also be addressed, followed by recommendations for further research on English language learners and testing.

Discussion

First and foremost, this research documents 11 English language learners'
experiences and understandings of high stakes testing - what others have described
tangentially, and often as second hand reports. I believe strongly that the voices of
students lend credibility, validity and urgency to several points, namely the need for
transparency and flexibility in our testing policies, the impact of perceived stakes and the
resulting ownership for all aspects of a test, and a consideration of unintended
consequences previously unexplored in testing literature. Underscoring these discussions
is a critical need to recognize English language learners as unique beings – despite their
holistic designation when disaggregated as a subgroup. Individual characteristics
(moderating variables) such as English proficiency level, time in the US, and former
schooling contribute to the ways in which they make sense of a phenomenon, and the
degree to which they engaged in sense-making as participants in this study.

Several inconsistencies in testing policy emerged from this study. The first, as
described by participants, is the requirement that all ELL, regardless of proficiency level,
take the TEAE every year. This requirement, coupled with the need to take both the
reading and writing portions every year (even if a student has achieved a proficient score
in one of those domains in the past) is an issue of test construction that can hopefully be
remedied relatively quickly. As participants suggested, a consideration of a student's
proficiency level is warranted in terms of both the language proficiency test (an adaptive
test by proficiency level), and the diploma test (an exemption based on proficiency, not

167

time). All ELL are not alike; one-size fits all policies do not work with ELL as a subgroup any more than they do with "all students". There is recognizable futility in requiring students to take a high stakes test of reading, math or writing in English when English proficiency is significantly below grade level. Additionally, perhaps the degree to which ELL are in need of multiple retakes of the diploma test could be avoided if students were allowed to first take the test when their facility with English was at an appropriate level, and not simply as a matter of course based on their grade level. At a minimum, students should be given a choice in the matter. More drastically, perhaps flexibility with regards to graduation paths is warranted. That is not to say that ELL should not be held to the same level of rigor as other students. But recognition of the accommodations and/or supports ELL must have to achieve at that same level is needed. If ELL do well in a class where linguistic modifications are made (and expected) yet the content standard is met, is it fair to judge them (with such high consequences) on a test that is not linguistically modified? I would argue that diploma tests are a contradiction of standards based education; they indicate a lack of trust in teachers and schools to teach to and evaluate against standards. Otherwise, wouldn't passing a class be good enough? If teachers and schools are not using rigorous standards to guide programming and instruction, to what extent are they complicit in perpetuating a false sense of accomplishment among learners when grades are inflated?

If in fact diploma tests are here to stay (as they most likely are), then perhaps other flexibility can be considered. For example, a process of appeal would allow a

student like Maria (who had taken BSTs multiple times and had come within 6 points of passing) to demonstrate her knowledge through a combination of alternative measures – measures that would document learning over time, be tied directly to instruction and therefore allow for the removal of the strong affective variable often created by a high stakes test.

Under the new accountability system in Minnesota, the tests used to demonstrate academic proficiency in reading and math will now also have graduation components. ELL in the US one year or less will be allowed to substitute the TEAE for reading and an accommodated test for math in terms of reports of adequate yearly progress, but how this will impact the graduation requirement components of the test remains to be seen. Minnesota is essentially combining two purposes into one test with the new system.

When NCLB tests and diploma tests were separate, the findings of this research demonstrate inconsistencies in ownership on the part of the school depending on for whom the stakes are highest. While tests required under Title I (more so) and Title III (less so) of NCLB are often referred to as high stakes in the media and research, in reality, they are high stakes for schools and districts more so than for students. This distinction is important when it comes to test implementation. As the data for this study show, there is less of an incentive for schools to take ownership of students' experiences with graduation test retakes under the current system (failure to stop announcements), as well as little ownership of the TEAE. Neither of these tests have direct consequences to the school. Language proficiency is reported by school *district* via Annual Measureable

Achievement Objectives, and attention to BST pass rates, if reported, gets lost in the shadows of Title I tests and reports of Adequate Yearly Progress.

The lack of systemic ownership for these tests has led to a variety of unintended consequences. For both tests, multiple retakes resulted in students doubting their academic and linguistic ability, which caused some participants to feel stigmatized. For the TEAE, a review of the test design could ameliorate many of the complaints of the test being below proficiency level. For both tests, the amount of instruction time lost to testing is another indication of sketchy ownership for these tests at the school level, and is an unintended consequence of the testing policies. The extent to which tests have value and are valued is contradicted if they become a burden to a select group of students. If participation in statewide tests is valued, shouldn't the time spent testing be respected and valued as well? In other words, students need not bear the burden of making up work due to mandated absences. As Crawford notes: "How can we ensure that the 'solution' does not exacerbate the problem" (2004, p. 1)? As students feel increasingly burdened and stigmatized by tests as a result of systemic factors, there is a risk of schools perpetuating inequity – particularly when different exams (and the students who take them) elicit different degrees of attention, consideration and respect.

An additional burden (and therefore unintended consequence) for English language learners is the extent to which they become "knowledge brokers" of testing information to their parents. The responsibility given students to explain test scores and consequences to their parents places them in an awkward and often frustrating position

when they are often uncertain of the details themselves. Families of English language learners often lack the cultural and linguistic capital necessary to advocate for fair and accurate information regarding accountability and graduation tests, let alone to advocate for fair testing conditions. Overall, there is a need for teachers, administrators and policy makers to provide opportunities for discussions of testing policies and for those discussions to be premised on clarity and transparency. As demonstrated in chapter four, students relish an opportunity to engage in discussions about testing policy, and want to know why and for what purpose they take the tests they do. Without such knowledge, misperceptions persist, and can lead to otherwise avoidable feelings of anger and frustration. Such dialogue may clear up misperceptions, and at the same time may bring into focus inconsistencies and unintended consequences of the policy, as was the case of this research.

As a primary implication of this research, I return to an exploration of sense-making and the importance of engaging in opportunities to do so with students. Given that the basic premise upon which different approaches to educational policy and practice rest is trust– trust that students will engage with tests, be invested in the accountability movement, cooperate…- how can we then move trust towards a more reciprocal relationship, creating spaces for students' voices to move out of the "pockets of progressivism" (Cook-Sather, 2002, p. 4) where they've traditionally been allowed to flourish, and into the mainstream of how we construct and enact policy? As Silva and Rubin (2003) point out, the role of students as active agents in, "…enacting or resisting

171

reforms remains vastly underestimated" (p. 2). My data reveal the complex, multidimensional nature of sense-making. The meaning participants ascribe to testing policies are often influenced by pre-existing knowledge and experiences, as well as those formed during the implementation of a test, during the research process, or by a variety of outside factors. Therefore, the process of coming to an understanding is iterative, non-linear, and complicated.

To understand the process, it behooves teachers and researchers to "become students just as students need to become teachers in order for education to become reciprocal and empowering for both" (Nieto, 1994, p. 399). Allowing students to speak reveals insights and complexities that go beyond those allowed by the "well-intentioned but largely destructive, scramble for performance" (Fielding, 2001, p. 102) that define our current testing policies. Considering young people's lives and experiences "as a lens through which to examine both the policy and the instructional dilemmas that now surround the education of immigrant children in this country" (Valdez, 2001, p. 2) is one tactic. Creating a two-way mirror is another. While researching adolescents can be arduous, dialogic research can be as equally rewarding and educative for the researcher as for the researched. As Michael Fullan asks, "what would happen if we treated the student as someone whose opinion mattered in the introduction and implementation of reform in schools?" (as cited in Corbett & Wilson, 1995, p. 15). Students need to be included in the *process* of change and not simply thought of as passive beneficiaries of reform or change, as they play a critical role in the success or failure of any reform.

172

Student voice research shares a basic premise: "…interpretations and representations of students' experiences, perspectives, and identities must be informed by students themselves, and furthermore, that the process of being informed is a matter both of acting and of being acted upon" (Cook-Sather, A., 2007, p. 829). Alison Cook-Sather, drawing on constructivist approaches to teaching in which teachers "follow the learner", suggests that researchers do this as well (2007, p. 856). Citing Honig and Hatch (2004), Luis et al. (2005) argue that policy coherence is created through practices of 'collective sensemaking' (p. 179). Ultimately, students are a voice in reforms to which they are subjected (Mitra, 2007; Rudduck, 2007; Thiessen, 2007; Yonezawa & Jones, 2007). The challenge lies in moving across the spectrum of student voice research, from listening to students to allowing them to become empowered.

Will creating spaces for student to make sense of their experiences lead to agency? I would argue that in the case of this research, it did, though minimally. An outcome of this research immediately following the data collection was an opportunity for participants to meet with an employee from the Minnesota Department of Education's LEP assessment division and share some of their concerns and suggestions regarding statewide testing. There is more to be done, however. A researcher's responsibility for data related to educational policy does not end with the submission of a paper. Fine et al. (2000) describe moving across the researcher-researched hyphen and into the application of their data by testifying at hearings and getting involved with the projects that were part of their research. Dennis Thiessen, in his introduction to the International Handbook of

173

Student Experience in Elementary and Secondary School describes an increase in action,

participatory, and activist-based research designs:

> Whether working for or with students who are different from the mainstream, who struggle in the margins, or who are part of the majority, researchers in these action-oriented projects are taking a stand on the kind of experiences students should have in classrooms and schools and, either within the cycles of fieldwork or as a planned follow up to the data gathering phase, joining with research participants to improve the voice, choice, and engagement of students. (2007, p. 55)

It is my intent to share the results of this research with other key stakeholders, and to

build on these data through future studies in conjunction with students. Primarily, the

following key suggestions for stakeholders emerged from the data.

Suggestions for Further Research and /or Policy Changes

My experiences suggest students are eager to have their voices heard. The

students in my study were high school aged English language learners currently enrolled

in school. This oft under-represented group of students and their opinions warrant further

research experiences, as do other students whose voices are seldom heard. Borrowing

from Magnuson (2003), I suggest action for various stakeholders as a step towards

addressing the findings presented in this study.

For policy makers

- Redesign the language proficiency test, and allow it to substitute for both the Title I test of reading (a federal decision), as well as the diploma test for reading and writing (a state level decision). Doing so would streamline

174

testing for ELL by reducing the number of required tests by 2 in attempt to alleviate testing fatigue.

- While redesigning the TEAE, consider either computer adapted leveling, or some measure of a locator test. Participants in this study indicated a real desire to be tested at their level and not be forced to start significantly below that level (as is the case with the current test). Doing so would alleviate some of the stigma students reported when required to test below their proficiency level.

For school personnel

- Make testing policies and practices transparent. Explain the rationale to students, and allow them to ask questions and participate in school level decisions when feasible.
- Carefully consider the consequences of statewide tests at all times, especially when those consequences bear high stakes for students.
- Attend to the instructional needs of all students, all the time.

For students

- Ask questions. Do not assume policies and resulting practice are static. Teachers, administrators and policy makers want to do the right thing, but need guidance and information from all stakeholders.
- Recognize the value of your experiences and insight. Many adults do not fully understand the complex nature of learning in a new language.

175

- Advocate for change, and share your knowledge with all who will listen.

Limitations

As with any research there are limitations in the design, implementation of and the sharing of results that warrant consideration. Primarily, it is important to acknowledge that this research did not occur in a vacuum, but rather in the context of adolescent lives in high school, and amidst all of the crazy, vibrant energy that entails. In general, qualitative, interpretive research allows for the requisite attention to be paid to the nuances and layers of students' experiences, as well as time for the ongoing dialogue that is necessary to build trust and reciprocity of understanding. Time was, to some extent, a limiting fact for the current study. Given the relatively short window for testing and the importance of discussing tests chronologically close to their implementation, inquiry groups took place at the end of the school year, setting a clear boundary for the research in terms of the number of times groups could realistically meet.

An attempt at efficiency was a factor in determining the criteria for participant selection. Students with lower levels of English proficiency were not included in this study, nor were students who had dropped out of school. Fielding notes how "the pressures of needing rapid results (of student voice research work) may lead us to listen most readily to voices that make immediate sense", emphasizing the importance of taking time with "anomalous" voices, from which we may actually learn more (p. 305). An attempt was made to explore and elaborate on anomalies in this research, though many

176

other voices were not present at the table. And what of the voices not included in the research—those who are truly silenced? How does excluding kids no longer in school contribute to the "silencing" of their voices (Fine, 1992), or those who didn't sign up for this study? The absence of other opinions in this data, though shared indirectly with reports of what participants' friends thought, are important to recognize and are part of the overall complexity of student experience with testing.

Yonezawa and Jones describe additional limitations in using students' voices to evaluate secondary school reform:

> One of the reasons that federal and state governments, districts, and schools turn to standardized-testing data to measure themselves is because quantitative scores provide quick summaries of how schools are doing year to year. Quantitative scores also provide an easy comparison of schools or districts with others…student voice data, admittedly, reflects the perspectives and experiences of only a tiny subset of students…interview and respondent bias may also warp the results (2007, pp. 707-708).

Participants in this study, by the very act of agreeing to participate, bring some bias to the results, as did I. While trying to remain true to the voices of students and to avoid leading them to certain answers, inevitably this may have occurred.

As a framework, sense-making was limiting in some respects until I was able to break out of the 'tiers' and theorize a more holistic approach. Though useful overall, I borrowed from research that stopped at teachers in terms of policy implementation agency. More needs to be done within the field of educational policy research to examine students as agents.

177

Finally, while a relatively new and under-explored method of data collection, I believe inquiry groups hold great promise for student voice research. This form of data collection was new for me, as well. There are undoubtedly ways in which my attempt at inquiry groups can be improved, though I believe so strongly in the possibilities for engaging students in this approach that I intend to employ it again.

Conclusion

Despite demonstrated flaws in No Child Left Behind, as a policy it does have the "potential to raise expectations and combat a "*pobrecito"* syndrome in classrooms – in which ELL students are cast as needing sympathy more than high expectations" (Black, 2006, p. 213). The building of teacher and administrator capacity to work with the diversity of students in schools and classrooms needs to become a priority (Olsen, 1997; Roderick & Engel, 2001). Any given attempt at reform such as those expected as a result of accountability testing stands little chance of success unless a *transformation* occurs; "simplistic technical responses to what are essentially moral and political dilemmas" are inadequate (Nieto, 1994), and the inclusion of students in the process is imperative.

Testing policies with ties to No Child Left Behind or its predecessors have been described as assimilationist in their exclusive focus on English (Menken, 2006) and for their "performance-focused discourses" (Black, 2006, p. 216). A critical examination of the intersection of such policies with the key stakeholders – students - is warranted. A further exploration of the intersection of such policies with the educational experiences of

178

those students learning English must be done expeditiously and shared widely. Perhaps the voices of students are just what policy makers need in order to form a balanced interpretation of the intentions and effects of testing policy.

As Thiessen and Cook-Sather remind us, "repositioning students in schools and in research on schools is a political stance" (2007, p. 583). So is re-examining the very notion of "all students" that underscores current testing policy in the US. Such statements are ubiquitous in educational policy in order for them to become politically popular, and unpacking such statements by examining the diversity of voice threatens to undermine support for policies (Spillane, 1998). At the same time we must remember that we cannot learn "once and for all…we must continually relearn to listen – in every context, with each group of students, and with each individual student" (Cook-Sather, 2002, p. 11). Even within the realm of student voice research, it is imperative that individual voices not become lost in the collective. To honor those voices, we must create spaces for dialogue, invite participation, and most importantly, listen.

References

Abedi, J., Hofstetter, C., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Educational Research, 74(1),* 1-28.

Abernathy, S. (2007). *No child left behind and the public schools.* Ann Arbor: University of Michigan Press.

Abrams, L. M. (2004). *Teachers' views on high-stakes testing: Implications for the classroom.* (Policy Brief). Arizona State University, Tempe. Education Policy Studies Laboratory.

Amrein, A., & Berliner, D. (2002). *An analysis of some unintended and negative consequences of high-stakes testing* (No. EPSL-0211-125-EPRU). Arizona State University, Tempe. Education Policy Studies Laboratory.

Anderson, M. E. (2004). Intended and unintended consequences of statewide testing for ESL curriculum and instruction. (Doctoral Dissertation, University of Minnesota, Twin Cities).

Anderson, M., & Stone, K. (2007). *Teachers, students, and statewide testing perceptions.* Presentation at the Fifth International Conference on Language Teacher Education, Minneapolis, MN

Anderson, M.E., Liu, K.K., Swierzbin, B., Thurlow, M.L., & Bielinski, J. (2000). *Bilingual accommodations for limited English proficient students on statewide reading tests: Phase 2* (No. 31 Minnesota Report). Minneapolis, MN: University of Minnesota, national Center on Educational Outcomes.

Balderrama, M., & Diaz-Rico, L. (2005). *Teacher performance expectations for educating English learners.* Allyn & Bacon.

Banks, J. A. (1998). The lives and values of researchers: Implications for educating citizens in a multicultural society. *Educational Researcher, 27*(7), 4-17.

Batt, L., Kim, J., & Sunderman, G. (2005). *Limited English proficient students: Increased accountability under NCLB.* (Policy Brief). Cambridge, MA: Harvard Civil Rights Project.

Bernstein, S. N. (2005). Teaching and learning in Texas: Accountability testing, language, race, and place. *Journal of Basic Writing, 23*(1), 4-24.

Black, W. R. (2006). Constructing accountability performance for English language learner students: An unfinished journey toward language minority rights. *Educational Policy, 20*(1), 197-224.

Blalock, G., & Haswell, R. (2003). Student views of TAAS. Available at http://comppile.tamucc.edu/TAAS/index.html

Center on Educational Policy. (2006). *From the capitol to the classroom. Year 4 of the No Child Left Behind Act*. Washington, D.C.: Center on Educational Policy.

Chabran, M. (2003). Listening to talk from and about students on accountability. In M. Carnoy, R. Elmore, & L. Siskin (Eds.). *The new accountability* (pp. 129-145). New York: RoutledgeFalmer.

Coltrane, B. (2002). *English language learners and high-stakes tests: An overview of the issues.* (ERIC Digest No. EDOFL0207). U.S.; District of Columbia: ERIC Clearinghouse on Languages and Linguistics.

Cook-Sather, A. (2002). Authorizing students' perspectives: Toward trust, dialogue, and change in education. *Educational Researcher, 31*(4), 3-14.

Cook-Sather, A. (2006). Sound, presence, and power: "Student voice" in educational research and reform. *Curriculum Inquiry,* 36:4 pp. 360-390.

Cook-Sather, A. (2007). Translating researchers: Re-imagining the work of investigating students' experiences in school. In D. Thiessen and A. Cook-Sather (Eds.). *International Handbook of Student Experience in Elementary and Secondary School(* pp. 829-871)*.* Dordrecht, The Netherlands: Springer.

Corbett, H. D., & Wilson, B. L. (1995). Make a difference with, not for, students: A plea to researchers and reformers. *Educational Researcher, 24*, 12-17.

Crawford, J. (2004). No Child Left Behind: Misguided approach to school accountability for English language learners. *Forum on Ideas to Improve the NCBL Accountability Provisions for Students with Disabilities and English Language Learners.* Washington, D.C.: Center on Education Policy.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed method approaches* (2nd Ed.). Thousand Oaks, Calif.: Sage Publications.

Education for Limited English Proficient Students Act, Minnesota Statutes, §§ 124D.58 - 124D.64

Feuerverger, G., & Richards, E. (2007). Finding their way: Immigrant and refugee students in a Toronto high school. In D. Thiessen and A. Cook-Sather (Eds.). *International Handbook of Student Experience in Elementary and Secondary School,(* 555-575)*. Dordrecht, The Netherlands: Springer.

Fielding, M. (2001). Beyond the rhetoric of student voice: new departures or new constraints in the transformation of 21st century schooling? *FORUM,* 43(2), 100-110.

Fielding, M. (2004). Transformative approaches to student voice: Theoretical underpinnings, recalcitrant realities. *British Educational Research Journal, 30*(2), 295-311.

Fine, M. (1992). Silencing and nurturing voice in an improbable context: Urban adolescents in public schools. In Fine, M. (Ed.), *Disruptive voices: The possibilities of feminist research* (pp. 115-138). Ann Arbor: University of Michigan Press.

Flores, B., & Clark, E. (2003). Texas voices speak out about high-stakes testing: Preservice teachers, teachers, and students. [Electronic version]. *Current Issues in Education, 6*(3).

Freeman and Freeman (2003). *Helping middle and high school age English language learners achieve academic success.* Available online at: http://www.uc.edu/njrp/pdfs/Freeman.pdf

Fullan, M. (1991). *The new meaning of educational change*. New York: Teachers College Press.

Gottlieb, M. (2003). *Large-scale assessment of English language learners. addressing educational accountability in K-12 settings* (TESOL Professional Papers #6). Alexandria, VA: Teachers of English to Speakers of Other Languages.

Gottlieb, M. (2006). *Assessing English language learners. Bridges from language proficiency to academic achievement.* Thousand Oaks, CA: Corwin Press.

Graue, M. E., & Walsh, D. J. (1998). *Studying children in context: Theories, methods, and ethics*. Thousand Oaks, Calif.: Sage Publications.

Haw, K. F. (1996). Exploring the educational experiences of Muslim girls: Tales told to tourists-- should the white researcher stay at home? *British Educational Research Journal, 22*(3), 319.

Holdwsorth, R. (2000, p. 358). Schools that create real roles of value for young people. *UNESCO International Prospect, 3*, 349-362

Jones, M., & Yonezawa, S. (2002). Student voice, cultural change: Using inquiry in school reform. *Equity & Excellence in Education, 35*(3), 245-254.

Kim, J., & Sunderman, G. L. (2004). *Does NCLB provide good choices for students in low-performing schools?* Cambridge, MA: The Harvard Civil Rights Project.

Kliebard, H. M. (1995). *The struggle for the American curriculum, 1893-1958 (2nd ed.).* New York: Routledge.

Koyama, J. P. (2004). Appropriating policy: Constructing positions for English language learners. *Bilingual Research Journal, 28*(3), 401-423.

LaCelle-Peterson, M. (2000). Choosing not to know. How assessment policies and practices obscure the education of language minority students. In A. Filer (Ed.), *Assessment: Social Practice and Social Product.* New York, NY: Routledge/Falmer.

LaCelle-Peterson, M., & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review,* 64(1), 55-75.

Lazarin, M. (2006). *Improving assessment and accountability for English language learners in the No Child Left Behind Act (Issue Brief No. 16).* Washington, D.C.: National Council of La Raza.

Lee, E., & Simon-Maeda, A. (2006). Racialized research identities in ESL/EFL research. *TESOL Quarterly*, 40(3), 573-594.

Leung, C., Harris, R., & Rampton, B. (2004). Living with inelegance in qualitative research on task-based learning. In Norton, B., & Toohey, K. (Eds.), *Critical pedagogies and language learning*. (pp. 242-267). Cambridge; New York: Cambridge University Press.

Lincoln, Y. S. (1995). In search of students' voices. *Theory into Practice, 34*, 88-93.

Louis, K. S., Febey, K., & Schroeder, R. (2005). State-mandated accountability in high schools: Teachers' interpretations of a new era. *Educational Evaluation & Policy Analysis, 27*(2), 177-204.

Macedo, D. P. (2006). *Literacies of power : What Americans are not allowed to know* (Expand ed.). Boulder, Colo.: Westview Press.

Magnuson, P. (2003). *The interplay between Minnesota's accountability system and adolescent English language learners with limited formal schooling.* (Doctoral Dissertation, University of Minnesota, 2003). *Digital Dissertations,* AAT 3092768.

McCall-Perez, Z. (1991). The Latina adolescent immigrant experience, participatory research: Listening to voices in a newcomer center. (Doctoral dissertation, University of San Francisco). *Dissertation Abstracts International, 54,* 11.

Menken, K. (2006). Teaching to the test: How No Child Left Behind impacts language policy, curriculum, and instruction for English language learners. *Bilingual Research Journal,* 30(2), 521-546.

Merriam, S. (1998). *Case study research in education: A qualitative approach.* Jossey-Bass.

Miles, S. (2004). *Student voices for school change.* Retrieved July 19, 2005 from http://www.alliance.brown.edu/stories/story0704.shtml

Miles, M., & Huberman, A. (1994). *Qualitative data analysis: an expanded sourcebook.* Thousand Oaks, CA: Sage Publications.

Mitra, D. L. (2004). The significance of students: Can increasing "student voice" in schools lead to gains in youth development? *Teachers College Record, 106*(4), 651-688.

Mitra, D. (2007). Student voice in school reform: from listening to leadership. In D. Thiessen and A. Cook-Sather (Eds.). *International Handbook of Student Experience in Elementary and Secondary School*(pp. 727-744). Dordrecht, The Netherlands: Springer.

Moraes, M. (1996). *Bilingual education : A dialogue with the Bakhtin circle*. Albany: State University of New York Press.

National Research Council. (1999). *Testing, teaching, and learning*. Washington, D.C.: National Academy Press.

184

Nichols, S., & Berliner, D. (2007). *Collateral damage. How high-stakes testing corrupts America's schools.* Cambridge, MA: Harvard Education Press.

National Clearinghouse for English Language Acquisition. (2006). *Resources About Assessment and Accountability for ELLs*. Washington, DC: Author. Available at http://www.ncela.gwu.edu/resabout/assessment/index.html

Nichols, S., Glass, G., & Berliner, D. (2005). High-stakes testing and student achievement: Problems for the No Child Left Behind Act. East Lansing, MI: Great Lakes Center for Education Research and Practice.

Nieto, S. (1994). Lessons from students on creating a chance to dream. *Harvard Educational Review, 64*, 392-426.

No Child Left Behind Act of 2001, 10 U.S.C. 6301 et seq. (2002) Retrieved July 10, 2006 from http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Oldfather, P. (1995). Songs "come back most to them": Students' experiences as researchers. *Theory into Practice, 34*, 131-137.

Oldfather, P. (1995). This issue {learning from student voices}. *Theory into Practice, 34*, 86-87.

Olsen, L. (1997). *Made in America: Immigrant students in our public schools*. New York: New Press: Distributed by W.W. Norton.

Patton, M. (2002). *Qualitative research and evaluation methods.* Thousand Oaks, CA: Sage Publications.

Peshkin, A. (1988). In search of subjectivity--one's own. *Educational Researcher, 17*, 17-21.

Pinar, W., Reynolds, W., Slattery, P., & Taubman, P. (2004). *Understanding curriculum.* New York: Peter Lang.

Quest, C., Liu, K., & Thurlow M. (1996). *Cambodian, Hmong, Lao, Spanish-speaking and Vietnamese parents and students speak out on Minnesota's Basic Standards Tests* (Minnesota Report No. 12). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved [8/30/06], from the World Wide Web: http://education.umn.edu/NCEO/OnlinePubs/MnReport12.html

Rivera, C., & Collum, E. (2006). *State assessment policy and practice for English language learners : A national perspective*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Rivera, C., & Vincent, C. (1997). High school graduation testing: Policies and practices in the assessment of English language learners. *Educational Assessment, 4*(4), 335-355.

Roberts, A., & Locke, S. (2001). Tending school: A forum on the experiences of refugee and immigrant students in the United States system. *Interchange, 32*(4), 375-393.

Roderick, M., & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low-achieving students to high-stakes testing. *Educational Evaluation & Policy Analysis, 23*(3), 197-227.

Roy, J., Custodio, B., Earley, K., New Levine, L., O'Loughlin, J., & Smallwood, B. (2007). *From understanding to using the preK-12 standards.* Presentation at the 41st Annual TESOL Convention, Seattle, Washington.

Rubin, B. (2003). On different tracks. Students living detracking reform at a diverse urban high school. In Rubin, B. C., & Silva, E. M. (Eds.), *Critical voices in school reform: Students living through change* (pp. 31-55). London; New York: RoutledgeFalmer.

Rudduck, J. (2007). Student voice, student engagement, and school reform. In D. Thiessen and A. Cook-Sather (Eds.). *International Handbook of Student Experience in Elementary and Secondary School(* pp. 587-610). Dordrecht, The Netherlands: Springer.

Ruiz-de-Velasco, J., & Fix, M. (2000). *Overlooked and underserved: Immigrant students in US secondary schools.* Washington, D.C.: Urban Institute.

Schmidt, M., & Datnow, A. (2005). Teachers' sense-making about comprehensive school reform: The influence of emotions. *Teaching and Teacher Education* 21, 949-965.

Schultz, K. (1997). Crossing boundaries in research and teacher education: Reflections of a white researcher in urban. *Qualitative Inquiry, 3*(4), 491.

Shohamy, E. G. (2001). *The power of tests: A critical perspective on the uses of language tests*. Harlow, England ; New York: Longman.

186

Silva, E. (2003). Struggling for inclusion. A case study of students as reform partners. In Rubin, B. C., & Silva, E. M. (Eds.), *Critical voices in school reform: Students living through change*. (pp. 11-30). London; New York: RoutledgeFalmer.

Silva, B., & Rubin, E. (2003). Introduction. Missing voices: listening to students' experiences with school reform. In Rubin, B. C., & Silva, E. M. (Eds.), *Critical voices in school reform: Students living through change*. (pp. 1-7). London; New York: RoutledgeFalmer.

Spillane, J. P. (2001). Challenging instruction for "all students": Policy, practitioners, and practice. *Yearbook (National Society for the Study of Education), pp.* 217-241.

Spillane, J., Reiser, B., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72(3),* 387-431.

Stake, R. (2000). Case studies. In Denzin, N. K., & Lincoln, Y. S. (Eds.), *The Handbook of Qualitative Research* (2nd ed.) (pp. 435-454). Thousand Oaks, Calif.: Sage Publications.

Stevenson, E., & Waltman, K. (2005, December). *The impact of NCLB on instructional changes: A consequential validity study.* Paper presented at the annual meeting for the Iowa Educational Research and Evaluation Association, Cedar Falls, IA.

Stone, K., & Anderson, M.E. (2006). [Focus groups with adolescent ELL regarding large-scale assessments]. Unpublished raw data.

Sunderman, G. L., Tracey, C. A., Kim, J., & Orfield, G. (2004). *Listening to teachers: Classroom realities and No Child Left Behind*. Cambridge, MA: The Harvard Civil Rights Project.

Sutton, M., & Levinson, B. A. (2001). *Policy as practice: Toward a comparative sociocultural analysis of educational policy*. Stamford, Conn.: Ablex Pub.

Taylor, G., Shepard, L., Kinner, F., & Rosenthal, J. (2002). *A survey of teachers' perspectives on high-stakes testing in Colorado: What gets taught, what gets lost.* (CSE Technical Report No. CSETR588). University of California, Los Angeles. Center for the Study of Evaluation, & National Center for Research on Evaluation, Standards, and Student Testing.

Teachers of English to Speakers of Other Languages. (2005). *Position paper on assessment and accountability of English language learners under the No Child Left Behind Act of 2001 (Public Law 107-110).* Alexandria, VA: TESOL.

Teachers of English to Speakers of Other Languages. (2006). *Statement of principles for the reauthorization of the Elementary and Secondary Education Act (ESEA).* Alexandria, VA: TESOL.

Thiessen, D. (2007). Researching student experiences in elementary and secondary school: An evolving field of study. In D. Thiessen and A. Cook-Sather (Eds.). *International Handbook of Student Experience in Elementary and Secondary School(* pp. 1-76*).* Dordrecht, The Netherlands: Springer.

Thompson, S., Thurlow, M., & Walz, L. (2000). *Student perspectives on the use of accommodations on large-scale assessments* (Minnesota Report No. 35). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved [8/31/06], from the World Wide Web: http://education.umn.edu/NCEO/OnlinePubs/MnReport35.html

Tillman, L. C. (2002). Culturally sensitive research approaches: An African-American perspective. *Educational Researcher, 31*(9), 3-12.

United States. Dept. of Education. Office of the Under Secretary, & United States. Office of Elementary and Secondary Education. (2002). *No child left behind : A desktop reference / prepared by the office of the under secretary*. Washington, D.C.: U.S. Dept. of Education, Office of the Under Secretary, Office of Elementary and Secondary Education.

United States. Department of Education. (2004). Secretary Paige's press release, 2/19/04. Retrieved from http://www.ed.gov/news/pressreleases/2004/02/02192004.html

United States. Office of English Language Acquisition, Language Enhancement, and Academic Achievement for Limited English Proficient Students. (2003). Part II. Final non-regulatory guidance on the Title III state formula grant program – standards, assessments and accountability. Washington, D.C: U.S. Dept. of Education.

Valdes, G. (2001). *Learning and not learning English: Latino students in American schools. Multicultural education series*. U.S.; New York: Teachers College Press.

Valenzuela, A. (1999). *Subtractive schooling: U.S.-Mexican youth and the politics of caring*. Albany: State University of New York Press.

Wheelock, A., Bebell, D., & Haney, W. (2000). Students self-portraits as test-takers: variations, contextual differences, and assumptions about motivation. *Teachers*

College Record. http://www.tcrecord.org ID Number: 10635, Date Accessed: 10/3/2006.

Wright, W. E. (2002). The effects of high stakes testing in an inner-city elementary school: The curriculum, the teachers, and the english language learners. [Electronic version]. *Current Issues in Education* [Electronic Version], *5*(5).

Wright, W. (2005). *Evolution of federal policy and implications of No Child Left Behind for language minority students* (No. EPSL-0501-101-LPRU). Arizona State University, Tempe. Education Policy Studies Laboratory.

Yonezawa, S. & Jones, M. (2006). Students' perspectives on tracking and detracking. *Theory into Practice 45(1), 15-23.*

Yonezawa, S., & Jones, M. (2007). Using students' voices to inform and evaluate secondary school reform. In D. Thiessen and A. Cook-Sather (Eds.). *International Handbook of Student Experience in Elementary and Secondary School(* pp. 681-710)*.* Dordrecht, The Netherlands: Springer.

Appendix A

Letter from Principal

Dear parents and guardians of English language learners:

I am writing to tell you about an opportunity for your child to talk about statewide tests. I have given a[district] teacher my permission to conduct a study of English language learners (Ells) at [Aldrich] High School. Her name is Karla Stone. Karla is an ESL teacher in the district and a graduate student at the University of Minnesota.

With your permission, Karla will meet with students in small groups and interview them about their opinions of tests. She may also interview your son or daughter individually. She will ask about the Test of Emerging Academic English (TEAE), and the Basic Standards Tests (BSTs). The interviews will take place during study halls, advisory periods, or immediately after school (depending on what is best for your son or daughter).

If your son or daughter participates, he/she will receive a $10.00 gift card to Target.

As principal of [Aldrich], I have also given Karla permission to review the transcripts (including test scores) of students who participate, which I will provide.

Please see the enclosed flyer and consent forms for more information.
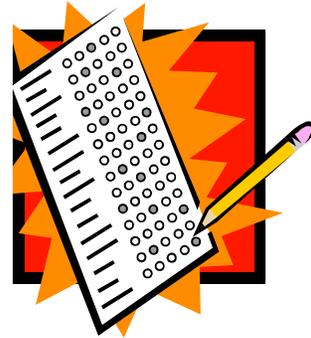
Sincerely,

David Dahl,
Principal,
[Aldrich] High School

Appendix B

Study Invitation

Attention

English

language

learners (ELL)!

- Are you an English language learner?

- Have you taken, or will you have to take a test this spring with a lot of others students in the state (BST, TEAE)?

- Would you like to share ***your opinions*** about tests?

If you answered 'yes' to these questions, consider being part of a study this semester. You are invited to share your opinions of tests in small groups and individually (by yourself).

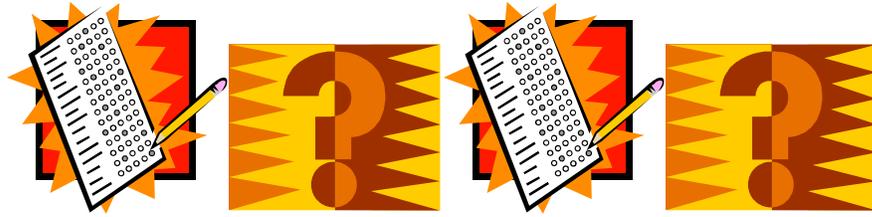If you participate you will receive a $10.00 Target gift card.

Please have your parents **read and sign the attached form** and mail it back to me. If you are 18, you may sign the form yourself.

I want to know what YOU think about tests!

Karla Stone, ESL teacher (on sabbatical) and graduate student, University of Minnesota (612) 626-9285 or ston0032@umn.edu

Appendix C

Questionnaire

Please complete the following sentences

1) The test I today (or recently) was called…

2) I took this test because…

3) The scores will be used to…..

4) Who else took this test?

5) Is the test important?
   Why or why not?

Please circle the best answer:

6) On this test, I tried…

really hard (my best effort)   pretty hard   a little bit   not much

7) When I took this test, I felt… (circle as many as you want)

   happy   sad   tired   anxious   bored   angry

   excited   confident   frustrated   hungry   sick   scared

Appendix D

Anticipatory Task 1

Name_____

1st language _____

What language do you speak most often at home? _____

Were you born in the U.S.? _____

   If not, when did you arrive in the U.S.? _____

What does T-E-A-E stand for?

What does B-S-T stand for?

What do you remember / think about the TEAE test?	What do you remember / think about the BST tests?

Appendix E

Anticipatory Task 2

Name

Sentence Frames

I think tests are

My friends think tests are

The people who make tests should

When the school gives tests, they should

Testing Analogies

Example: Winter is to cold, as summer is to heat.
                Winter : cold :: summer : heat

School is to _____ as testing is to _____.

The TEAE is to _____ as the BSTs are to _____.

Others:

194

Anticipatory Task 3

Name

You write the test!

1. **The TEAE test is_____**
    a.

    b.

    c.

2. **The BSTs are _____**
    a.

    b.

    c.

3. **English language learners should take tests when _____**

    a.

    b.

    c.

4. **Students take tests like the TEAE and BST because _____**
    a.

    b.

    c.

5. **All students _____ have to take tests.**
    a.

    b.

    c.

Others

195