Teachers as Critical Users of Assessment for Emerging Bilingual Students

Ashley Meghan Chrzanowski
University of Colorado Boulder, ashleychrz@gmail.com

Follow this and additional works at: https://scholar.colorado.edu/educ_gradetds
Part of the Bilingual, Multilingual, and Multicultural Education Commons, and the Educational Assessment, Evaluation, and Research Commons

Recommended Citation
Chrzanowski, Ashley Meghan, "Teachers as Critical Users of Assessment for Emerging Bilingual Students" (2015). School of Education Graduate Theses & Dissertations. 73.
https://scholar.colorado.edu/educ_gradetds/73

This Dissertation is brought to you for free and open access by School of Education at CU Scholar. It has been accepted for inclusion in School of Education Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
TEACHERS AS CRITICAL USERS OF ASSESSMENT FOR EMERGING BILINGUAL STUDENTS

by

ASHLEY CHRZANOWSKI

B.A., Scripps College, 2006
M.A., Hunter College, 2009

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado at Boulder in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
Department of Education
2015
The thesis entitled:
Teachers as Critical Users of Assessment for Emerging Bilingual Students

Written by Ashley Chrzanowski
Has been approved for the Department of Education, University of Colorado at Boulder

Guillermo Solano-Flores, Ph.D.

Kathy Escamilla, Ph.D.

Lucinda Soltero-González, Ph.D.

Kara Viesca, Ph.D.

Claudia Mills, Ph.D.

Date________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB Protocol # _14-0078_
The purpose of this exploratory study was to determine the extent to which teachers are critical users of assessment for emerging bilingual students and to identify which factors affect their critical use. A critical user of assessment is a teacher who engages in critical thinking and action. Critical thinking is defined as a practice that takes place within social context that develops the skill of reasonable, reflective thinking. In this work, the notion of critical thinking operationalizes Kahneman’s (2010) model of System I and II thinking. In this model, System II describes effortful mental activity, whereas System I describes routine thinking. The conceptual framework proposes ‘critical engagement’ as an additional component, describing activities based on critical thinking.

Participants included 82 teachers. 72 teachers were surveyed, six “focus” teachers were interviewed and observed, and four teachers were only interviewed. Half of the teachers had experience in a Culturally and Linguistically Diverse (CLD) educational program. The data sources were composed of interviews, classroom observations, and survey information. Data sources were organized and coded to identify the proportion of activities that represented the various components of critical thinking and action: System I, System II, and the Critical Engagement (CE) component. Narratives of the six focus teachers were developed based on the data. Trends in the survey data were identified.

Of the three components, results indicated that System II was used the least. Some teachers engaged in CE component activities more than others, showing a higher level of critical use of assessment. Characteristics of teachers who engaged in CE component activities
frequently were identified to create a profile of a critical user of assessment. CLD program experience proved to be inconclusive. School, district, and state level requirements emerged as a factor with a negative influence on teachers’ critical use of assessment. This evidence suggests that a strict focus on data driven instruction disempowers teachers to act critically in their practice. This highlights the need for providing teachers with the freedom and support to develop as critical users of assessment so they can, in turn, foster habits of critical thinking in their emerging bilingual students.
Acknowledgements

I appreciate Dr. Solano-Flores for your thoughtful guidance and unending support since the first step of this process. You have caused me to grow as an academic and as an individual. Dr. Ruiz-Primo, you have been a tremendous support as well. The two of you welcomed Mike and me with open arms into your lives, and we will never forget this.

This dissertation is dedicated to the emerging bilingual students with whom I was privileged to teach. You taught me more than I ever expected. I hope that you have access to the opportunities to which you aspire.

I want to thank all the incredible teachers who work hard every day to provide access for emerging bilingual students, often without any recognition. For the teachers who spend their free time and money making sure that their students have beds to sleep on, that their bellies are full, and that their minds are hungry for learning. For Patty, Myrna, and Mary: You were so supportive and generous. I will never forget that. For Mr. Llinas, Professor Odell, and Mr. Georgi: Thank you for believing in me.

I am eternally grateful for my husband, Mike. You have been my rock throughout this entire process. I most certainly would have never arrived here if it had not been for you. You have supported me in every way and have always done it lovingly. We have made many transitions over this period of time and you have made them go smoothly. Thank you for everything you have done and continue to do to make our family your priority. I am grateful for our daughter, Maya. Not a day goes by without me feeling blessed that you came into our lives. I am working to create a world you will be happy to live in.

I am grateful for my parents, Lori La Posta and Dave Kost. You have made sacrifices throughout your life so that Jeff and I could have the privilege of educational opportunity. Becoming a parent has made me begin to understand how just much you have given to us. I am appreciative of how you continue to care for us. Thank you. To my brother, Jeff, thank you for always lending a patient, listening ear and giving me your honest opinion.

I am thankful for the members of my dissertation committee, as you all provided helpful guidance in my educational journey. I am also thankful for all of my dear colleagues at the University, especially Karla. You rejuvenated my spirit and gave this program new meaning. Chelsey, Jackie, Cristin, Clara, Katy, Meg, Juli, Heidi, Jenn, Kristin, and Yingying: Thank you for being a shoulder to lean on!

I want to acknowledge the Educational Testing Service for generously providing funding for this project through the Small Grant for Doctoral Research in Second or Foreign Language Assessment award.
Table of Contents

Chapter 1: Introduction ........................................................................................................1
  Statement of the Problem .................................................................................................6
  Purpose Statement ...........................................................................................................9
  Research Questions .........................................................................................................10
  Significance ....................................................................................................................11

Chapter 2: Conceptual Framework ....................................................................................13
  Sociocultural Theory ......................................................................................................16
  Orientations to Language ...............................................................................................17
  Cognitive Theory ............................................................................................................18
  Critical Perspectives on Assessment with Emerging Bilingual Students ..................23
  Critical Thinking ...........................................................................................................24
  My Conception of Critical Thinking ............................................................................25
  Critical Pedagogy ...........................................................................................................26
  Distilled Conceptual Framework ....................................................................................27

Chapter 3: Review of Literature .........................................................................................35
  Section One ...................................................................................................................35
  Section Two ...................................................................................................................44
  Section Three ...............................................................................................................62

Chapter 4: Methodology ....................................................................................................72
  Strategy of Inquiry ..........................................................................................................72
  Participants and Design .................................................................................................73
  Researcher Role ..............................................................................................................77
  Instrument Development and Data Sources .................................................................78
  Recruitment and Data Collection Procedure ...............................................................84
Data Analysis.................................................................................................................89
Chapter 5: Findings.......................................................................................................98
Teacher Narratives.....................................................................................................98
   Adele.........................................................................................................................99
   Alma.........................................................................................................................101
   Valera.......................................................................................................................103
   Lucas.......................................................................................................................106
   Serena.......................................................................................................................108
   Claire.......................................................................................................................110
Survey Results............................................................................................................112
Research Question One, System II.............................................................................130
   Conclusion...............................................................................................................141
Research Question Two, CE component.....................................................................142
   Conclusion...............................................................................................................155
Research Question Three, System I...........................................................................156
   Conclusion...............................................................................................................179
Chapter 6: Discussion..................................................................................................183
   Research Question One..........................................................................................183
   Research Question Two..........................................................................................185
   Research Question Three........................................................................................185
   Profile of a Critical User of Assessment...............................................................188
   Reflections on the Profile of a Critical Assessment User.......................................189
   Comparing Profiles of Critical Users of Assessment.............................................191
   Implications for Policy.............................................................................................192
   Limitations...............................................................................................................197
List of Tables

Chapter 2
Table 1. Research Questions and Relevant Domains That Enable me to Address These Questions Effectively .......................... 16

Chapter 3
Table 2. Definitions of Critical Thinking. Taken From R. H. Johnson (1992) ...................... 47

Chapter 4
Table 3. Research Design ............................................................................................................. 76
Table 4. Interview Protocol 1 Items and Their Corresponding Critical Thinking Activities and Components ................................................................. 79
Table 5. Sample of the Preliminary Coding Method for the Classroom Observation Data .... 94

Chapter 5
Table 6. Correspondence Between the Range of Percentages of Component-Related Activities in Which Teachers Engaged and the Frequency Descriptor Used to Describe Them ............................... 98
Table 7. Response to “I receive the most helpful information about my ELLs' progress towards CONTENT AREA goals from: ” Percentages .......................................................... 113
Table 8. Response to “I receive the most helpful information about my ELLs' progress towards LANGUAGE PROFICIENCY goals from: ” Percentages .................................................. 114
Table 9. Response to “When it comes to designing and implementing high-quality assessment methods to measure how my ELLs are progressing with a unit of study, I feel: ” Percentages .................................................. 115
Table 10. Response to “When I receive helpful information about my ELLs' progress in a class, I MOSTLY: ” Percentages .......................................................... 117
Table 11. Response to “I MOSTLY think assessment is: ” Percentages .................................. 118
Table 12. Response to “In general, I create my own assessments” Percentages .......................... 124
Table 13. Response to “I change or modify parts of assessments that I do not create myself” Percentages ...................................................................................... 125
Table 14. Response to “My colleagues and I collaborate to plan our assessment methods” Percentages ...................................................................................... 126
Table 15. Percentage of Reported Component-Related Activities for Each Focus Teacher Based on Interview, Observation, and Combined Data Sources .................................................. 131
Table 16. Types of Modifications That Teachers Make to Their Assessments for Emerging Bilingual Students and Relative Frequency of These Modifications .................................................. 152
Table 17. Component-Related Activities of Focus Teachers Grouped by CLD Educational Experience in Percentages ...................................................................................... 176-177
List of Figures

Chapter 2

Figure 1. Broad view of perspectives relevant to critical use of assessment with emerging bilingual students............................................................................................................................14
Figure 2. Distilled and focused conceptual framework relevant to teachers’ critical use of assessment with emerging bilingual students.................................................................................................28
Figure 3. Linear progression of agency from System I to the CE component..........................................................................................................................32

Chapter 5

Figure 4. The proportion of component-related activities Adele employed in interview sessions..........................................................................................................................100
Figure 5. The proportion of component-related activity Adele employed in classroom observation periods.................................................................................................................................101
Figure 6. The proportion of component-related activities Alma employed in interview sessions..........................................................................................................................102
Figure 7. The proportion of component-related activity Alma employed in classroom observation periods.................................................................................................................................103
Figure 8. The proportion of component-related activity Valera employed in interview sessions..........................................................................................................................105
Figure 9. The proportion of component-related activity Valera employed in classroom observation periods.................................................................................................................................106
Figure 10. The proportion of component-related activity Lucas employed in interview sessions..........................................................................................................................107
Figure 11. The proportion of component-related activity Lucas employed in classroom observation periods.................................................................................................................................108
Figure 12. The proportion of component-related activity Serena employed in interview sessions..........................................................................................................................109
Figure 13. The proportion of component-related activity Serena employed in classroom observation periods.................................................................................................................................110
Figure 14. The proportion of component-related activity Claire employed in interview sessions..........................................................................................................................111
Figure 15. The proportion of component-related activity Claire employed in classroom observation periods.................................................................................................................................112
Figure 16. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to me, as a teacher” percentages.................................................................................................................................119
Figure 17. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to my students” percentages.................................................................................................................................120
Figure 18. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to my students’ parents” percentages.................................................................................................................................121
Figure 19. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to my principal and administration” percentages.................................................................................................................................122
Figure 20. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to my district” percentages.................................................................................................................................123
Figure 21. Coded responses from both CLD-endorsed and non-CLD-endorsed teachers to “If you could change anything about assessment used with ELLs (either large-scale assessments or classroom-based assessments), would you?” All respondents said yes and then were prompted to respond to: “What would you change, and why?”

Figure 22. Responses of teachers with and without CLD educational experience to the question: “What holds you back from making changes?”

Figure 23. Focus teachers’ relative engagement in System I, System II, and CE component activities from the interview data in percentages.

Figure 24. Focus teachers’ relative engagement in System I, System II, and CE component activities from the observation data in percentages.

Figure 25. Focus teachers’ relative engagement in System I, System II, and CE component activities from the combined interview and observation data.

Figure 26. Surveyed teachers’ coded responses to the question: “What holds you back from making changes?”

Figure 27. Word graphic displaying surveyed teachers’ comments regarding their lack of power or control with assessment for emerging bilingual students.

Figure 28. CLD and Non-CLD focus teachers’ relative engagement in System I, System II, and CE component activities from the combined interview and observation data in percentages.
Teachers as Critical Users of Assessment for Emerging Bilingual Students

_The important thing is to not stop questioning._
- Albert Einstein

This dissertation addresses the need for fair and valid assessment practices for emerging bilingual students and investigates how to better support teachers to use assessments more critically and thoughtfully. More specifically, this dissertation focuses on exploring the extent to which teachers are critical users of assessment for emerging bilingual students.

For the purposes of this work, an emerging bilingual student can be defined as a student who is acquiring English as a new language. This expression is synonymous with the terms, “English Language Learner,” “Non English Proficient,” and “Limited English Proficient,” which are also frequently used in educational contexts. I chose to refer to these students as “emerging bilinguals” because this term focuses on their developing linguistic abilities rather than their perceived deficiencies in the area of language. “Multilingual learners” is a similar term that can be used for this purpose.

In the context of this work, _assessment_ refers to the comprehensive process of eliciting evidence and drawing inferences about student knowledge and skills (Pellegrino, Chudowsky, & Glaser, 2001; Popham, 2000). Tests are a commonly employed tool to elicit evidence about student knowledge and skills. For this reason, the term “test” is often used interchangeably with “assessment” in the literature. However, a test is not the only medium through which one can elicit evidence about student understanding. Assessment also includes informal measures in which teachers elicit students’ understanding about the content in an unplanned manner. For example, a teacher can ask questions to students during a lesson and gather information based on their responses that help the teacher to infer a student’s level of understanding of a learning goal.
There are tests for different purposes (e.g., formative, diagnostic, interim, and summative assessment) and with various scale sizes, from classroom to large-scale assessment. In this work, I focus on assessment in a broad sense and specify when I refer only to a particular type of assessment. In the style of Cochran-Smith, Gleeson, and Mitchell (2010), I use the term *assessment* to refer to any assignment or activity that gives a teacher information about their students’ learning in this document.

Also for the purpose of this work, *critical use* involves two dimensions of analysis: critical thinking and critical action with regard to assessment for emerging bilinguals. In this work, critical thinking involves the use of reflective, reasoned thinking, and critical action involves agency and activism supported by critical thinking that promotes equity for emerging bilingual students.

**Challenges in the Assessment of Emerging Bilingual Students**

Emerging bilingual students are the most rapidly growing subset of the United States school-age population. Between the academic years 1997-98 and 2008-09, this segment of the student population grew by 51 percent, while the non-emerging bilingual segment only increased by seven percent (National Center for English Language Acquisition, cited by Samson & Collins, 2012). As of 2008, the emerging bilingual student population already totaled five million students, which was double the number from 15 years prior (National Education Association (NEA), 2008). The NEA predicts that, by 2025, almost one quarter of the total student population in the United States will be composed of emerging bilinguals (NEA, 2008). Given this demographic trend, it is of utmost importance to appropriately address the needs of emerging bilingual students in order to ensure that they have equitable access to opportunities during the
course of their education and beyond. Unfortunately, the current educational environment does not tend to foster equal opportunity for emerging bilingual student populations.

An important facet of this issue is the difference in performance between emerging bilingual students and non-emerging bilingual students on national large-scale assessments. Differences in achievement between emerging bilingual students and non-emerging bilingual students can be observed in the National Assessment of Educational Progress (NAEP), which tests representative samples of students in the United States in the fourth-grade and in the eighth-grade. In the 2011 NAEP administration in mathematics, emerging bilingual fourth- and eighth-grade students lagged behind their non-emerging bilingual peers by an average of 24 and 42 scale-scored points, respectively (NAEP, 2011a).

Performance differences are even more pronounced in reading, as the 2011 NAEP results indicate that fourth- and eighth-grade emerging bilingual students were, on average, 37 and 43 scale-scored points behind non-emerging bilingual students, respectively. These differences in performance are also reflected in assessments at both the district and state levels. For example, in 22 states, the first-time passing rates of emerging bilingual students on high school exit examinations in mathematics and reading were substantially lower than those of non-emerging bilingual students (Center on Education Policy, 2006).

The causes of these performance differences can be framed in a variety of ways, and the framing affects the ways in which people work to address these performance differences. In the view of Ladson-Billings, the achievement gap can be likened to an educational deficit, whereas the forgone school resources that should have been invested in low-income students can be likened to the educational debt (Ladson-Billings, 2006). We spend so much of our time attempting to reduce our educational deficit that we never manage to make a dent in the
educational debt that continually contributes to the deficit. The source of the problem is the educational debt, since many students do not have access to resources that would provide them with an equal opportunity to learn. This debt is not framed as purely financial by Ladson-Billings (2006), but also includes historical, sociopolitical, and moral debt. However, we spend our time as a nation working on reducing the educational deficit because more attention is paid to the current year deficit in terms of test score reporting. We are essentially treating the symptoms rather than the cause of the problem, thus we only scratch at the surface of a solution to the problem.

Instead of focusing on how to address the education debt, however, policy makers have focused on reducing the education deficit, or the “achievement gap.” This education deficit has been interpreted by some policy makers as primarily the result of ineffective teachers and teaching practices. Locally, Senate Bill 10-191 bases 50% of teacher evaluations on various student achievement and growth measures (Meyer, 2010; The Colorado Department of Education, 2010). The No Child Left Behind (NCLB) Act of 2001 was designed to address the “achievement gap” by increasing teachers’ and schools’ accountability for their students’ test scores (U.S. Department of Education, 2001). Some policy makers continue to attribute the “achievement gap” to teachers’ levels of accountability for their students’ test scores, as is evidenced by a large, recent competitive grant program called the Race to the Top initiative. In this competition for federal funds, states were given incentives to create high-stakes evaluation systems for teachers based on their students’ state test scores (U.S. Department of Education, 2009). The assumption underlying these federally supported priorities is that state-mandated testing programs yield valid and reliable evidence that can be used judiciously to reward effective teachers and to penalize those who are ineffective in the classroom.
While state-mandated tests are carefully developed due to the large amount of political and financial support allocated to these assessment programs, no single indicator is sufficient in making decisions about teaching effectiveness, instructional programs teachers use, or students’ language or content area learning. Campbell’s Law warns that the more a social indicator is relied upon to make decisions, the more likely it is to be subject to pressure from corruption and therefore distort the very processes it is designed to monitor (Campbell, 1976). These concerns may be especially serious if the law applies to large-scale assessment, as linguistically and culturally diverse populations and their teachers will be especially vulnerable to its effects when diversity is not properly addressed in testing programs.

Many issues have been identified that affect the validity of the large-scale assessment process for emerging bilingual students. As a sample, I will briefly discuss three of them: the test development process, the accommodations used to support emerging bilingual students, and the time frame for including newly identified emerging bilinguals in testing programs. First, emerging bilingual students are often underrepresented in the theoretical framework of assessment programs and during the process of test development (Solano-Flores, 2011); inferences are made about emerging bilingual students’ understanding of content based on results from tests developed for and with monolingual students.

Second, testing accommodations, or modifications made to tests to support students in showing their understanding about the tested content area, are not consistently administered to emerging bilingual students, and they vary drastically in their levels of effectiveness (Abedi, Hofstetter, & Lord, 2004; Rivera, Collum, Shafer Willner, & Ku Sia, Jr., 2006; Wright & Choi, 2005). In fact, some of the most commonly employed testing accommodations for emerging bilingual students, such as extra time and small group or individual testing administration, are
derived from accommodations designed for students with special needs and therefore fail to address emerging bilingual students’ linguistic differences (Abedi et al., 2004).

Third, in many cases, emerging bilingual students are included within content area testing programs after they have been registered in the school system for one academic year, which is an insufficient period of time to acquire the necessary levels of English language proficiency to access the items on content area tests. Research suggests that it takes approximately five to seven years for emerging bilingual students to develop English language proficiency (Hakuta, Butler, & Witt, 2000). Thus, while language proficiency is recognized as an important factor that challenges the validity of test scores for emerging bilingual students, it has not yet been effectively counteracted by testing accommodations.

**Statement of the Problem**

Many issues, from test development to test administration, challenge the assumption that state- and district- mandated assessment systems are valid and reliable for emerging bilingual students. Assessment policies and practices should change if these students are to be assessed fairly and validly. Teachers can be important agents in this change, as they are in the position to assess their students and observe the direct effects of assessment systems and policies within their classrooms. Unfortunately, literature is silent regarding the extent to which practitioners think about or question assumptions about testing programs, their ability to measure culturally and linguistically diverse students’ knowledge, and the appropriateness of using measures of student academic achievement as indicators of teacher effectiveness.

There are other factors relevant to understanding why teachers may not contest issues pertaining to assessment of emerging bilingual students. Teachers may not feel adequately prepared by their formal and informal experiences with assessment and related issues such as
validity, reliability, and utility. Although teachers spend a minimum of one third of their efforts in the classroom engaged in activities related to assessment (Stiggins, 1991), pre-service teacher education programs do not tend to address assessment in much depth. In fact, most states do not even require an assessment course in order to gain initial teacher certification (O’Sullivan & Chalnick, 1991; cited by Quilter & Gallini, 2000). Only one half of teacher preparation programs require undergraduate pre-service teachers to enroll in a measurement course (Jett & Schafer, 1992). Other teacher preparation programs embed the concept of assessment within other classes, often at the expense of being privy to an instructor with specific measurement expertise (Schafer, 1993). Thus, teachers are unlikely to have the level of assessment literacy necessary to confidently engage in critical thinking or critical action with regard to assessment.

Nguyen-Le (2010) asserts that “teacher education programs need to provide opportunities for teachers to develop critical thinking skills on issues of language and culture and connect these skills to the classroom” (p. 106). Yet the majority of teacher education students and teacher educators are white, monolingual English speaking, and with little academic experience or training in diversity (Keiser, 2005; Nieto, 2000). As of the 2007-08 academic year, 83% of public school teachers identified as non-Hispanic whites, while only 54% of students in public schools in the fall of 2009 identified as non-Hispanic whites (U.S. Department of Education and the National Center for Education Statistics, 2012). This lack of diversity among America’s teaching force contrasts with the large proportion of ethnic and racial minorities, which now comprise the majority of children born in the United States—50.4 percent of the total population (United States Census Bureau, 2012). The mismatch between the profile of the average child and a typical teacher within the public education system speaks to the importance of preparing
teachers of culturally and linguistically diverse students to be able to use culturally responsive assessment practices to help students develop their language abilities.

A premise guiding this dissertation is that, to properly serve their emerging bilingual students, teachers need to be able to navigate the domains of assessment and culturally and linguistically diverse education. In addition, teachers need to think and act critically in response to issues of assessment with these students. This work views teachers’ critical thinking about assessment as essential for the struggle for social justice in education. I contend that without a critical lens, assessment is only a tool for reproducing inequities for students from culturally and linguistically diverse backgrounds. Unfortunately, to date there is no body of research on how teachers think critically about the kinds of decisions that they are able to make regarding assessment for emerging bilingual students. The extent to which teachers critically consider issues of language and culture within their own assessment methods is also largely unknown. If the American public values the implementation of assessment practices that evaluate emerging bilingual students fairly, then teachers should have an active role in ensuring that this takes place.

Since the reauthorization of the Elementary and Secondary Education (ESEA) Act by President George W. Bush in 2001 - the NCLB Act, large-scale testing has taken a more prominent role in the evaluation of schools, teachers, and students. Yet, while assessment is a vital tool for practitioners to tailor their instruction to the needs of their students, particularly the linguistic needs of their emerging bilingual students, teachers may not view assessment as a powerful part of their classroom practice. Since many school-related decisions are based heavily on high-stakes testing scores, it is not surprising that many teachers feel as though assessment is an external process that does not belong to them.
Solano-Flores and Soltero-González (2011) posit that this could be the result of several factors. One issue is that, because of its association with pressure from accountability, teachers may perceive assessment as a tool for control instead of a resource to guide their instruction. Second, it can be difficult to use state-mandated tests as a tool for improving instruction, since test results are often delivered when school is no longer in session. Instead, most teachers receive these results when they no longer have the group of students on which scores are reported and they can no longer use them to inform their instruction. This can have a devastating and disempowering effect on practitioners. The extent to which assessment is meaningful to teaching also appears to play a critical role in the way teachers perceive it as belonging to their realm of activities. Students benefit when teachers and schools create a culture where learning is the focal point and teachers use “tests worth teaching to,” which Shepard (2000a) describes as being dichotomous to a culture of testing (p. 9). If an assessment does not contribute to the teaching and learning process, then it does not benefit students or teachers.

**Purpose Statement**

This study examines whether teachers have a critical awareness of cultural and linguistic issues involved in assessment with emerging bilingual learners, and the extent to which they are agents of improvement of assessment practices and use assessment methods and tools that fairly account for cultural and linguistic differences. The inquiry into if and how teachers of emerging bilingual students are critical users of assessment is a unique question that has not yet been studied empirically. In attempting to support teachers in becoming critical assessment users, Solano-Flores and Soltero-González (2011) emphasize the importance of appropriating assessment as a meaningful part of their teaching practice. They stress the need for a culture of assessment that promotes teachers as critical users of assessment to benefit their culturally and
linguistically diverse students. No empirical study has been conducted to assess the degree to which teachers already engage in these practices.

This dissertation addresses the need for teachers to be able to think critically and speak out as professionals if and when externally imposed assessments are inappropriate for their students, and to be otherwise autonomous in integrating valid classroom assessment practices into their instruction. Teachers need to have both the skills to think critically and the will to act upon their critical thinking in order to carry out these practices.

This dissertation also attends to the fact that, even when teachers have the ability to think critically about assessment, there may be factors that prevent them from acting on their conclusions. It is important to identify these limiting factors and how they can be addressed. If teachers are not able to express their concerns about assessment practices to school administrators and test developers, the opinions of the people who have the most direct view of how assessment policies and programs manifest themselves in the classroom will be disregarded. No educational innovation will live up to its potential if the innovators do not hear feedback from those who implement it.

**Research Questions**

The three main research questions that guide this study are:

1. To what extent do teachers of emerging bilingual students think critically about assessment?

2. To what extent do teachers of emerging bilingual students use assessment critically?

3. Which factors influence teachers’ critical thinking about and critical use of assessment for emerging bilingual students?
Significance

This study develops new knowledge on teachers’ skills in assessment for emerging bilinguals that will be useful to multiple audiences. This study questions the degree to which teachers recognize the elements that can contribute to assessment practices that are incongruous for emerging bilingual students and whether they take action to improve or limit the use of inappropriate classroom assessment and large-scale assessment practices. The study sheds light on the ways in which teachers that are critical users of assessment recognize and respond to assessments that do not serve emerging bilingual students properly. Ultimately, the study is intended to benefit teachers and therefore aims to contribute to empowering them as professionals in their practice and to encourage them to be critical users of assessment with emerging bilingual students. This study intends to contribute to support teachers in developing a critical consciousness surrounding appropriate assessment practices for emerging bilingual students.

Policy makers can benefit from this study, as it allows identification of the kinds of policy decisions that would support teachers in being critical users of assessment. The work provides actionable recommendations that policy makers can apply at the district, state, and federal levels to institute assessment practices and promote assessment views that directly attend to the needs of emerging bilingual students.

Finally, this study establishes the basis for a research agenda aimed at developing teachers’ critical thinking about and use of assessment with emerging bilingual students. This research agenda can be fostered through the participation of teacher educators and professional development leaders. These professionals may find value from this study and use it to guide future teachers in understanding what it means to be a critical user of assessment with emerging
bilingual students and how to bring this critical consciousness to their own practice. Based on evidence from this dissertation, leaders of professional development for teachers can also encourage in-service practitioners to bring a critical lens to their current assessment practices.

Raising awareness about how to promote equitable assessment practices for emerging bilingual students is essential, given that assessment is the primary tool used to make major decisions about students such as promotion, retention, and graduation. In the struggle for justice for emerging bilingual students in the U.S. education system, our nation must focus on the measures with which we evaluate students from different linguistic and cultural backgrounds.
Chapter 2
Conceptual Framework

The term, “critical user of assessment” was first used in Solano-Flores and Soltero-González’s (2011) paper on meaningful assessment in linguistically diverse classrooms. Although they do not provide a definition of a critical user of assessment, they allude to the idea that critical assessment users are practitioners who view assessment as a meaningful part of their practice, and who are active participants in the process of assessment. They stress that this development will benefit teachers as well as their emerging bilingual students.

I propose that there are two dimensions which characterize teachers as critical users of assessment. First, teachers who are critical assessment users question their own assumptions and analyze their assessment practices to determine if the practices they implement are equitable for their emerging bilingual students. They judge whether the inferences they make from their assessment practices are justified and whether they contribute to the teaching and learning process. Second, teachers who are critical assessment users take action that is based on their critical thinking. Teachers who resolve issues that arise when they evaluate their assessment practices serve to limit inequitable assessment practices for emerging bilingual students, unjustified inferences about students based on assessment practices, or assessment practices that do not contribute to teaching and learning processes.

In this section, I describe a conceptual framework that offers a lens through which to explore teachers’ critical thinking about and critical use of assessment with emerging bilingual students. Figure 1 depicts the conceptual perspectives relevant to critical thinking about and critical use of assessment with emerging bilingual students. I view critical pedagogy as a concept that can be housed within sociocultural theory because the concepts are compatible with each other. Sociocultural theory focuses on culturally-relevant and meaningful social interactions as a
driving force to incite learning (Rogoff, 2003). This focus is well-suited to accompany the idea in critical pedagogy that one can use this learning through dialogue to aide people in the process of their liberation from oppression through a combination of reflection and action (Freire, 1970). While sociocultural theory and critical pedagogy are centered on the way people learn and transform through social interaction, cognitive theory focuses heavily on how we make sense of the world as an individual (Kahneman, 2010). Cognitive theory draws on many cognitively based processes to form ideas about how we can make reasonable judgments based on the information we receive about the world (Ennis, 1987).

![Figure 1. Broad view of perspectives relevant to critical use of assessment with emerging bilingual students.](image)

Although sociocultural and cognitive theory are as different conceptually from each other as critical pedagogy and critical thinking are different in their traditions, they can both enrich our perspective without being mutually exclusive, and they can even inform each other at times. The pairing of sociocultural and cognitive theory has already been accomplished: Shepard (2000a) wrote about merging sociocultural and cognitive theory to gain insight from both theories in what she called “social-constructivist” theory. Additionally, a bridge exists that connects the
traditions of critical pedagogy and critical thinking. This bridge was constructed by authors who include elements of both traditions in their writing (e.g. Atkinson, 1997; Brookfield, 1997; Cummins, 1989; Curry, 1999; Giroux, 1994).

The elements of critical thinking and critical pedagogy play a direct, valuable, and distinct role in understanding of how teachers are critical users of assessment. I identify elements of critical thinking and critical pedagogy that are both essential to contextualizing how teachers think and use assessment critically for emerging bilingual students. These critical traditions lay the groundwork for understanding how thought and action work together to enable teachers to be agents of change for more equitable and justified assessment practices for emerging bilingual students. The tradition of critical thinking discipline highlights the importance of mostly cognitively-based thinking processes while the tradition of critical pedagogy primarily informs socially-based action, both of which I believe must be present in order to identify a practitioner as a critical user of assessment (Ennis, 1987; Freire, 1970).

A final, yet essential, piece of my conceptual framework is the concept of one’s orientations to language. This concept centers my work, since a teacher’s orientations to language color their thoughts and actions in particular ways. Some teachers may be aware of their orientations to language and therefore have the potential to maintain a critical consciousness about their orientations. Some teachers may not be conscious of their language orientations and how they impact their assessment of students. Regardless of their level of awareness, teachers’ orientations to language diversity invariably impact the way they use assessment with emerging bilingual students. All of the aforementioned concepts shape the way that I interpret how teachers think about and use assessment with emerging bilingual students.
In Table 1, I show the relevant domains that enable me to address each research question of this study effectively. While some domains are more relevant than others in addressing a specific research question, all of the domains work together to inform my study as a whole. In the section that follows, I discuss how each domain plays a role in my study, how the discipline informs my understanding of how teachers use assessment with emerging bilingual students, and how it renders this work more comprehensive.

Table 1

Research Questions and Relevant Domains That Enable me to Address These Questions Effectively

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Supporting domains, from most to least relevant</th>
<th>Main concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To what extent do teachers of emerging bilingual students think critically about assessment?</td>
<td>Critical thinking, orientations to language, cognitive theory, sociocultural theory, critical pedagogy</td>
<td>Critical thinking regarding assessment</td>
</tr>
<tr>
<td>2. To what extent do teachers of emerging bilingual students use assessment critically?</td>
<td>Critical pedagogy, orientations to language, sociocultural theory, cognitive theory, critical thinking</td>
<td>Critical action regarding assessment</td>
</tr>
<tr>
<td>3. Which factors influence teachers’ critical thinking about and critical use of assessment for emerging bilingual students?</td>
<td>All domains equally inform this question</td>
<td>Influencing factors</td>
</tr>
</tbody>
</table>

Sociocultural Theory

This work is situated in the tradition of sociocultural theory, which conceptualizes learning as a primarily social process. Shepard (2000) posits that understandings which develop in the mind are culturally and socially determined, as put forth by Vygotsky (1978).

Furthermore, students learn through an active process that is enhanced by meaningful social interaction, which stands in opposition to the traditional passive role of students within the
classroom (Cummins, 1989; Wiliam, 2007). Students work together in classroom communities of practice to move forward in their learning of content area knowledge. Through this process, students gradually move from a peripheral role in the community of practice to a more integral one. Rogoff (1995) describes this transition as moving from a role of apprenticeship to that of guided participation, and then finally to adopting the role of participatory appropriation, where the student engages as a full and leading member of the community of practice. Teachers may also form communities of practice when they collaborate with each other in order to plan, reflect upon, and improve their practice, including assessment methodology. Communities of practice provide teachers with the opportunity to learn from their colleagues as well as to contribute new understandings to the group (Lave & Wenger, 1991; Wenger, 1998).

**Orientations to Language**

A sociocultural perspective of language then, in my view, considers different cultural and linguistic backgrounds as a resource for student learning since students can learn from each other’s different experiences. Proponents of sociocultural theory would encourage students to use their array of experiences to help enrich their learning and to help them think critically about the content area knowledge they study. This view corresponds to the “language as a resource” orientation put forth by Ruiz (1988), also as used by Escamilla, Chavez, and Vigil (2005). Teachers who adopt the “language as a resource” orientation believe that bilingualism is advantageous to learning and development. When considering assessment methodology, these teachers’ assessments would reflect the understanding that emerging bilingual students are learning content area knowledge and language simultaneously, and that their English language abilities may vary depending on the context of a particular assessment or interaction (Grosjean, 1985).
Conversely, if an educator characterizes cultural and linguistic differences as negative traits that students should not maintain, she has a “language as a problem” orientation. If an educator believes that one should maintain their native language, yet they consider language and culture to be factors that hold students back academically, they would possess a “language as a right” orientation. The “language as a problem” and “language as a right” orientations both espouse a deficit perspective of emerging bilingual students to varying degrees, though a teacher with the “language as a problem” view has a stronger propensity to attempt to assimilate students with linguistic and cultural differences. A practitioner with the “language as a right” view tends to problematize differences between emerging and non-emerging bilingual students within a school setting, while still promoting the maintenance of these differences in their home environment. Teachers with a “language as a right” or a “language as a problem” orientation may not be as aware of the way that issues of language might interact with a teacher’s assessment of a student’s content area knowledge.

**Cognitive Theory**

While learning is undoubtedly shaped by social interaction, culture, and community, the cognitive processes that take place in the mind also play an important role in the learning process. Shepard (2000) posits that sociocultural, cognitive, and constructivist theories all have important insights to contribute to the understanding of teaching and learning processes, and she developed a theory to reflect this belief: social-constructivist theory. She claims that understandings from cognitive theory can work in tandem with socio-cultural theory to create a more coherent lens for understanding curriculum, learning theory, and assessment. The conceptual framework for this study builds upon the idea that people learn through socially
supported interactions, which then enables the mind to use existing knowledge to interpret new knowledge to create meaning (Bransford, Brown, & Cocking, 1999).

The aspects of cognitive theory that contribute most to the current study include Kahneman’s (2010) concept of thinking fast and slow, the idea that prior knowledge and cultural understandings influence learning, the concept that metacognitive skills are important to learning and thinking, and the idea that cognitive performance is dependent upon one’s dispositions and personal identity. These concepts provide a helpful lens to understand and contextualize how one’s mental processes influence their perceptions and their critical thinking.

The concept within the field of cognitive theory that is most relevant to this study is captured in Kahneman’s (2010) book, *Thinking, fast and slow*. In this work, Kahneman suggests that the mind has two distinct ways of interpreting information. He expands upon Keith Stanovich and Richard Wests’ (2000) original labeling of two systems in the mind, System I and System II. System I functions relatively automatically and rapidly, with little effort and control. An example of System I at work would include a person’s ability to correctly interpret someone’s facial expression to mean that they are upset or the ability to understand simple sentences. Conversely, System II can be characterized by mental activities that require effort and control, such as computation, concentration, agency, and choice. An instance in which System II is activated could be when a person performs a long division mathematics problem or focuses on one person’s voice in a noisy crowd. System II conducts the kind of effortful action necessary for critical thinking in general, and specifically with regard to assessment with emerging bilingual students. Conversely, System I can function with relative automaticity, and may be used by a teacher to make copies of a unit test of a textbook for their class. Kahneman (2010) argues that
these two systems work together very efficiently by minimizing effort and optimizing performance for our mental processing activities.

However, there are disadvantages to the efficiency obtained when these systems work together. For example, Kahneman (2010) asserts that System I can be biased, or the mind may be “lazy” and fail to employ System II where necessary, thus allowing the mind to be deceived by certain scenarios (p. 31). In this way, Kahneman lays the foundation for researchers to understand the contexts in which critical thinking processes are more or less likely to occur and why. This study focuses on the extent to which teachers employ critical thinking skills, using System II processes, when engaged in assessment activities for their emerging bilingual students. Additionally, this study attempts to identify the factors that influence teachers’ use of critical thinking in assessment. The concept of System II in cognitive processing and the contexts in which it is or is not activated, may prove useful for understanding the factors that influence a teacher’s critical thinking in assessment with emerging bilingual students.

The idea that prior knowledge and cultural understandings influence learning constitute a second key concept from cognitive theory (Shepard, 2000b). This concept is widely understood in pedagogical practice for children, but is not as often considered with respect to adult learning (Echevarria, Vogt, & Short, 2008). Prior knowledge and cultural understandings shape the way people interpret subsequent knowledge and develop their own understandings of new phenomena. The more knowledge and cultural understandings with which people are familiar, the easier it is for people to make nuanced meaning from new information. This informs my study because awareness of the variation in teachers’ prior knowledge and cultural understandings helps me to understand differences across teachers and the way their critical thinking skills have been developed. For example, if a teacher views language as a resource, this
is a cultural understanding that shapes the way they interpret and act upon new information about emerging bilingual students. Likewise, if a teacher has extensive background in the area of assessment, this knowledge shapes the way in which they create and use assessments in their classroom. Therefore, relevant prior knowledge and cultural understandings have a large impact on the critical thinking skills of teachers with regard to assessment of emerging bilingual students.

The concept that metacognitive skills are important to learning and thinking is relevant to the current study because these skills help people to monitor their own critical thinking processes (Brown, 1994; Shepard, 2000b). Metacognition refers to thinking that involves active control over the cognitive processes central to learning (Livingston, 1997). Pintrich and De Groot (1990) showed that self-regulation, as carried out through metacognitive strategies for planning, monitoring, and modifying cognition, was an important predictor of academic performance. Shepard (2000b) noted that metacognitive skills can be acquired through processes that are socially mediated, identical to the way in which cognitive skills are learned.

An example of successful use of metacognition in educational programs includes reciprocal teaching (Palinscar & Brown, 1984), which provides students with socially scaffolded practice using four metacognitive strategies: predicting, developing questions, summarizing, and clarifying. In a different study on metacognition, Sternberg (1992) identified “executive processes,” which are sequential components of the metacognitive self-monitoring process that include (1) recognizing the existence of an issue, (2) determining the nature of the problem, (3) choosing a set of lower-order processes to solve the issue, (4) creating a strategy to combine the processes, (5) choosing a mental model of the problem, (6) distributing one’s cognitive
resources, (7) monitoring one’s problem solving techniques when they are used, and assessing one’s own problem solving once complete (Shepard, 2000b).

Sternberg’s (1992) executive processes inform this work, as these processes may act as a guide with regard to the steps that teachers take in order to think critically. For example, if a teacher does not perform the first executive process in the sequence, which is recognizing the existence of a problem, then the teacher will not be able to think critically about problems when they exist. Thus, Sternberg’s (1992) executive processes may help in the process of assessing various components of metacognition, which facilitate critical thinking.

Additionally, an important concept from cognitive theory is that cognitive performance is influenced by one’s characteristics and identity (Bandura, 1997; Dweck, 1999; Shepard, 2000b; Yin et al., 2008). External evaluations of a student’s academic performance, particularly those that are person-centered instead of behavior-centered, are highly likely to impact the student’s perception of themselves as a learner. The expectations of a teacher, as communicated either verbally or through written assessment results, may have a significant influence on a student’s ability to achieve their best in the classroom. This helps to explain how critical use of assessment on the part of teachers is so important for students, since invalid assessment results may influence a student’s perception of themselves as a learner. This effect can be compounded over time, and can have a significant effect on the trajectory of a learner throughout their educational experience. Additionally, a teacher’s sense of self-efficacy influences their inclination to engage in critical thinking and to take action as a result of their thinking. If teachers do not feel confident to engage in critical thinking activities, then it is more likely that they will not do so (Bandura, 1995). Teachers’ identities as expressed through their beliefs and actions may impact their critical use of assessment as well. If teachers do not initially take action on issues they deem
unethical, then the cognitive dissonance between their belief on the issue and their corresponding inaction could lead them to change their belief about the issue over time (Rasinski, Geers, & Czopp, 2013). Thus, self-efficacy and identity of teachers also play a role in the process of critical use of assessment by teachers.

**Critical Perspectives on Assessment with Emerging Bilingual Students**

I argue that teachers should employ a critical lens in the way they conceptualize the notion of learning in the classroom, particularly regarding the role of language in promoting learning through their classroom assessment practices. In order to include these critical notions in my framework, I sought the resources of the critical thinking and critical pedagogy domains. While these two traditions are distinct and sometimes even stand in opposition to each other, these two theories can work together to bring a more comprehensive lens to the study I plan to conduct. The authors who take up the concept of critical thinking with the inclusion of the social dimension include Dewey (1933), Paul (1992), Benesch (1993), Brookfield (1997), Atkinson (1997), and Curry (1999). Scholars who focus on critical pedagogy specifically include Freire (1970), Aronowitz and Giroux (1985), Giroux (1994), Marx (1845), Apple (2008), McLaren (1994), Kincheloe (2007), and Cummins (1989). These scholars believe that we learn through social processes to some extent, and therefore they embed this belief into their theories of how people learn to think critically. In the next section I discuss critical thinking and its relevance to the study. Then I detail critical pedagogy and how, together, the disciplines afford a more comprehensive understanding of teachers’ critical use of assessment with emerging bilingual students.
Critical Thinking

Critical thinking is relevant to this study, and education in general, as it highlights the kind of thinking that is desirable for students to engage in and to transfer to their own lives after their formal educational experience is complete (Ennis, 1987; Ravitch, 2010). In order for students to be able to engage in these kinds of thinking processes, teachers must model the qualities of good critical thinkers. Even if teachers explicitly teach critical thinking, if they don’t themselves use the skills they teach, students tend to learn from what teachers do instead of what they say (Dewey, 1933). Critical thinking is not simply possessing skills that can be taught; it is also knowing when and how to use them appropriately to make reasonable, reflective decisions about what to do or believe (Ennis, 1987). Teachers need to be good role models for critical thinking in their own lives in order for these qualities to carry over to the next generation. If educational experts ignore the need to develop long-term strategies for nurturing the development of teachers’ own critical powers, as Paul (1992) asserts, then an emphasis on critical thinking for students could easily turn into a form of social indoctrination itself.

One of the largest conceptual challenges of studying, teaching, and assessing critical thinking is the lack of clarity about the qualities that critical thinking entails across the field of critical thinking experts. This is rather ironic, since “clarity” is typically one of the qualities that are offered as a component of critical thinking (Ennis, 1987); however, experts have not yet been able to come to a clear consensus about this practice. Since even experts on the topic of critical thinking struggle to clearly define the components that comprise critical thinking, it is difficult to know when one is truly studying, teaching, or assessing it.

Additionally, the nature of critical thinking is extremely broad, and is generally not specific towards issues of assessment or of emerging bilingual students. There is no body of
literature focused on teachers’ critical thinking with regard to their assessment practices for emerging bilingual students specifically. The majority of critical thinking research is not written in a domain-specific manner, and some critical thinking experts argue specifically for the generalizability of critical thinking (Siegel, 1992). However, some experts believe that critical thinking is domain-specific and non-generalizable (McPeck, 1992). This rift demonstrates how critical thinking, as a field, is both unclear and extremely broad, since even the experts fail to arrive at a consensus of whether or not critical thinking is a general skill that applies to all subject areas or if it is particular to specific domains. Thus, in the following section, I conceptualize critical thinking from my own perspective, informed by the aforementioned theories of language and learning that I espouse. The remainder of this work employs the following conception of critical thinking so as to promote clarity with regard to the meaning of this terminology.

**My Conception of Critical Thinking**

My understanding of critical thinking is positioned largely from within the sociocultural perspective; it reflects a greater focus on the social context than most other definitions set forth by critical thinking experts. In my view, individual cognition plays an important role in critical thinking. However, individual cognition is also incessantly influenced by the communities in which we take part; through apprenticeship, guided participation, and participatory appropriation (Rogoff, 1995). I view critical thinking as a social practice rooted in culture (Atkinson, 1997). Therefore, the social context of critical thinking should be included in its definition to acknowledge the community of practice that it can create. Critical thinking, then, can be considered to be a practice that takes place within a specific social context that develops the skill
of reasonable, reflective thinking. I offer this conception of critical thinking in the hopes that work based on this notion can bridge the traditions of cognitive and sociocultural theory.

**Critical Pedagogy**

Critical pedagogy ensures that activism is linked to the notion of critical thinking. Though there is some resistance in the field of critical pedagogy to identifying the founding fathers of the movement, due to the implied structure of patriarchy which critical pedagogy hopes to counteract, it is important to understand the history behind the tradition (Breuing, 2011; Lather, 1998). Critical pedagogy stems primarily from the work of Paulo Freire, who wrote in the context of an impoverished northeastern Brazil in the 1960s. He emphasized pedagogy of the critical theory that began in the Frankfurt School of Germany.

With the first publication of Freire’s book, *Pedagogy of the Oppressed*, in 1967, he inspired a movement that became widely popular across North and South America. Freire was committed to using indigenous epistemology and knowledge as tools to promote the empowerment of and justice for indigenous people and those who understand indigenous epistemology (Kincheloe, 2007). The purpose of this pedagogy is to educate people to become “critical agents who actively question and negotiate the relationships between theory and practice” (Giroux, 2007, p. 1). Freire wrote that in order to be liberated from oppression, one must reflect and act upon the world in order to transform it (Freire, 1970). Thus, critical pedagogy demands that its’ supporters not only reflect critically upon these relationships, but also that they take action based on their critical thinking. Critical pedagogy also states that a central purpose for the valuing of activism is to promote democracy among an actively critical society. This critical lens includes a consideration of the dynamics of power inequities and why they exist.
Michael Apple (2008) discusses how it might look when teachers adopt critical pedagogy in their own lives. He explains that it is necessary for teachers to be able to reposition themselves in order to have a more complete understanding of social justice and the role of education in the struggle for social justice. Repositioning entails having the knowledge and critical social understanding that allows teachers to view the world through the eyes of those who are less privileged. Apple (2008) warns that the attempt to foster critical characteristics in teachers by promoting critical inquiry could actually be counterproductive if teachers are not able to reposition themselves from the viewpoint of their students. Apple’s argument suggests potential steps for subsequent action if this is the case. The current study; however, seeks to investigate if and how teachers are able to engage in critical inquiry within the context of assessment of emerging bilingual students.

**Distilled Conceptual Framework**

While Figure 1 provided a broad view of the theories that inform my understanding of teachers’ critical use of assessment with emerging bilingual students, Figure 2 depicts a more distilled, focused perspective of the elements that are involved in critical assessment use. The boxes on the far left of the figure show the various structures within which teachers operate in schools. These structures influence teachers’ thinking and actions, which are diagrammed into three sequential components: System I, System II, and the optional component of Critical Engagement (CE). Each component of teachers’ thinking and action includes progressively more agency, beginning from System I, to System II, and finally to the Critical Engagement (CE) component. From System II, teachers may engage in activities that have an impact on social, structural factors. In the active component, called Critical Engagement (CE), teachers may also act to incite change that then influences the social, structural factors. In multiple ways, then,
there is a relationship between the structures that exist in the teaching profession and the thinking and action systems of teachers.

![Diagram](image)

**Figure 2.** Distilled and focused conceptual framework relevant to teachers’ critical use of assessment with emerging bilingual students.

The structures shown on the left hand side are factors which influence how teachers socialize through work on a regular basis. The concept of learning a profession or a discipline as
a socialization process is taken from the sociocultural tradition (e.g. Moses, 2001). Though Moses describes learning as a process of socialization only in the context of students, here the notion is extended to teachers. Shleppegrell (2010) also discusses how a current, prominent topic in mathematics research is the idea that students are “apprenticed into particular ways of doing mathematics” (p. 74). This echoes Rogoff’s theory of the three planes of sociocultural activity as well as Lave and Wenger’s theory about how students learn within communities of practice. Lave and Wenger (1991) and Rogoff (1995) did not limit their theory to students in classrooms. In my view, learning is a process of social participation, not only with regard to the way in which students learn in classrooms, but also with regard to the way in which adults are acculturated into the teaching profession. Teachers are apprentices, quite literally, during their student teaching experience. Teachers also continue to learn through communities of practice with their colleagues, their principals, and professional development session leaders.

Within the boxes depicting structure, then, there are several factors that have a large impact on teachers. First, the school institutions all contain defined hierarchies, with traditions and rules that shape the ways in which teachers interact with authority figures and with their colleagues. Standards are of particular influence in the current era of standards based reform, when assessment and evaluation systems are designed around sets of standards for both students and teachers. The media also has an impact on how teachers are socialized into the profession, since these forms of communication influence the manner in which many people, including teachers, develop their ideology about the role education should play in children’s lives and their expectations about how teachers should serve students. Thus, the structures in which teachers play a part have a large impact on the activities in which teachers engage in System I, which are shown to the right of the structures. The socialization process is depicted using unidirectional
arrows, which indicate that people are influenced by structures when engaged in System I activities.

System I and System II are concepts derived from Kahneman’s (2010) aforementioned cognitive conceptual framework. Kahneman (2010) does not identify any additional components or factors in his model. In this work, I propose adding an optional component to the model, called the component of Critical Engagement (CE). These diagrammed thinking and action components also originate in Dewey’s (1933) depiction of reflective - as opposed to routine - thinking and action. Dewey distinguished between routine action, guided by tradition and authority, and reflective action, driven by active and persistent consideration (Zeichner & Liston, 1996). Kahneman’s System I has many similarities to Dewey’s described ‘routine action’, and they are both incorporated in the current System I component.

The thinking and action categorized in System I include cognitive activities that can be conducted with relative automaticity, as they do not involve thinking that is as systematic or rigorous. System I activities can be observable or unobservable, since they are mostly activities that involve adjusting to one’s community both mentally and physically. For example, in order to comply with new nationally administered assessments such as Assessing Comprehension and Communication in English State-to-State for English Language Learners (“ACCESS for ELLs®”) to evaluate emerging bilingual students’ proficiency in English, one does not need to engage in critical thinking as much as one needs to simply administer the test to students. The test administration procedure is very well defined and little agency is required to administer it. One can carry out skills like this in System I with a certain degree of passivity, aside from the socialization that takes place between a person and the structures around them. For instance, teachers may undergo training to learn how to administer the test. Beginning to compare or
contrast the new “ACCESS for ELLs®” test with other English language proficiency assessments, however, indicates that a person is beginning to think more actively and systematically about them.

As one moves from left to right in Figure 2 from System I to System II and optionally to the CE component, the solid arrows indicate that agency increases progressively. Figure 3 shows this progression of agency. Across all of the thinking and action components, there is an inverse relationship between the level of structural influence and the level of agency that each component entails. The level of structural influence is highest in System I, while the level of agency is lowest.

Since agency can have several distinct connotations across (and even within) disciplines, it is necessary to define how I am using the term, “agency.” In this dissertation, the meaning of agency is taken from Miller and McIntyre (2012) and Richert (1992), who state that agency refers to an individual or a group of people who believe that their voice matters in this world, and that agency serves as the connection between reflection and action. Some theorists view agency as a loop that feeds back to influence structure (Marx, 1845). Others posit that people use their own agency to reproduce the current structural system for the future generation (Bourdieu, 1977). Although it is possible to have a scenario with a structure with free agents, typically I believe that increased structural influences pose limits on individual and group agency. Thus, it follows that the level of agency grows when thinking and action become less tied to the influence of structural factors.
Figure 3. Linear progression of agency from System I to the CE component.

The activities that are included in System II involve more agency and critical thinking than System I activities, and they are less influenced by structural factors than those in System I. System II activities involve mostly unobservable skills, since they mainly occur as habits of the mind. System II allows someone to analyze the knowledge one has of the world around them and then make judgments or conclusions based on their analysis. These System II activities help people to better reflect upon the mental or physical activities in which they might engage in System I.

Lastly, the CE component characterizes activities that are least influenced by structural factors and which most involve one’s own agency. This component could, with more research, potentially act as an extension of Kahneman’s Systems I and II, as a component of critical engagement was not included in his model. The CE component contains activities that are mostly observable, such as organizing a group for a particular cause. For example, an educator might organize a group of fellow practitioners to collaborate in preparing their students for the new
“ACCESS for ELLs®” exam, by collectively thinking about ways to relate the new content that will be tested to the experiences of the students in their community. These activities might progress from (1) initially gaining familiarity about the structure of the new “ACCESS for ELLs®” test, then (2) comparing and contrasting the format and content of the new test with the previous one, and finally (3) collaborating with colleagues to prepare their students for the test and relate new content to their students more effectively. It is helpful to take this action after going through the progression in this framework, which begins with System I (gaining knowledge about the test) and proceeds to System II (analyzing the test), before arriving at the CE component (organizing the group to prepare students). However, arriving at each step is optional.

The CE component is the space in which paradigm shifts may be supported by people when they become familiar with new theories, analyze and test them against other dominant theories, and then choose to support a new theory (Kuhn, 1996). Though Kuhn only used the term “paradigm shift” to account for revolutions of theory in the scientific community, theorists in the humanities have borrowed the principle. I also capitalize on this principle in order to provide an analogy as to how paradigms might shift in education as well. Thus, an arrow connects System II, and optionally the CE component, to the boxes of structural factors in order to demonstrate that teachers may have an effect on structural factors if they engage in System II or CE component activities and use their agency.

Additionally, it must be made clear that this conceptual framework does not only pertain to the way in which individuals think and act. This framework for critical thinking applies to groups of people as well as to individuals. The way in which people organize and define themselves determines if they can be considered to think and act within this framework as
individuals or as a group. Thus, this framework is flexible in that it can encompass this progression of agency in both settings.
Chapter 3
Review of the Literature

Overview

As of yet, there is no literature that examines the extent to which teachers of emerging bilingual students are critical users of assessment. However, there is a limited research base that specifically focuses on teachers’ perspectives on assessment with emerging bilingual students. Despite the meager amount of literature on the topic, this chapter assembles a comprehensive review on the subject of teachers’ critical use of assessment with emerging bilingual students by drawing on various related areas of research. This literature review is separated into three distinct sections. The first section addresses the topic of teachers’ perceptions of assessment with emerging bilingual students through three themes: (1) authors who explicitly state their theoretical perspectives on language and learning, (2) an author who centers her work on the issue of fairness in testing as expressed through a focus on consequential validity, and finally, (3) authors who focus their research on policy-oriented issues related to federal, state, or local policies that drive large-scale and classroom assessment practices. The second section reviews extant literature on critical thinking and critical pedagogy, focusing on how the two traditions are distinct from one another and how the two disciplines might inform each other. The third section reviews both the standards for teachers’ assessment skills and the standards for the pedagogy of emerging bilingual students.

Section One: Teachers’ Perceptions of Assessment with Emerging Bilingual Students

Although there are many studies that pertain to assessment or emerging bilingual students or teachers’ perceptions and perspectives on their practice, there are some works that discuss all three of these topics. This literature review was performed using the following procedure: I searched for articles on Google scholar, the Education Resources Information Center (ERIC)
server, and the Elton B. Stephens Company (EBSCO) Host server using several search terms. These search terms included: teacher perceptions, teacher perspectives, teacher beliefs, teacher views, teacher conceptions, English Language Learners, Bilingual, English as a Second Language, and assessment. I used combinations of search terms that always included at least one term to refer to emerging bilingual students and assessment in my searches. Additionally, I used the selected articles presented in this section to search Google scholar for all of the articles that cited these selected articles and for those that were deemed to be ‘relevant’ to the selected articles according to Google scholar.

**Theme One: Explicit Theoretical Perspectives on Language and Learning**

Rueda and Garcia (1996) used a constructivist perspective to categorize teachers’ perspectives on assessment with emerging bilingual students. They investigated the beliefs and practices of teachers of emerging bilingual students with respect to literacy, bilingualism, and assessment. They had three distinct groups of 18 teachers as participants, each with different types of training: bilingual credentialed teachers, bilingual waivered teachers, and special education teachers of emerging bilingual students. The researchers conducted a two- to three-hour semi-structured interview with each teacher using a unique protocol to gather data on teachers’ beliefs in the areas relevant to the study. They organized teachers’ perspectives along a continuum for each category: from a transmission-oriented to a constructivist approach to instruction, from subtractive to additive views of bilingualism, and from discrete-skills to a constructivist approach to reading assessment. They developed a coding system for each category, and averaged an interviewee’s ratings over the course of an interview to obtain an average relative position of each interviewee on each of the three continua.
The findings revealed that, with regard to the bilingualism category, special education teachers were more likely to have negative, subtractive views of using native language in the classroom. Strikingly, however, no group of teachers all shared an additive view of bilingualism, not even the bilingual certified teachers. Likewise, the authors found that no group approached reading instruction in an entirely constructivist manner. With respect to views of assessment, Rueda and Garcia (1996) found that, again, no group of teachers was clustered at the constructivist-oriented end of the continuum. Rueda and Garcia (1996) noticed several trends that differentiated teachers with constructivist practices from the practices of teachers who had a discrete-skills orientation. They noticed that constructivists had ongoing assessments throughout the day and the unit, instead of scheduling a specific day and time to test students’ knowledge. They also found that students have more influence in the activities that take place in constructivists’ classes, and that students play more of a role in the assessment process in these classrooms as well. Finally, they noticed that all teachers expressed pressure from their administration for their students to learn English quickly, which detracted from even the bilingual teachers’ ability to focus on developing their students’ bi-literacy.

Rueda and Garcia (1996) argued that teachers’ beliefs about this were mediated by features of the context of the schools’ attitude towards learning English. The authors understood that teachers may very well recognize the discrepancy between their “imperfect applications” of theories, such as second language acquisition theory, but that teachers recognize them as reasonable adaptations to their environment. The authors hypothesized that, since there were group differences between bilingual and special education teachers about reading and assessment practices, prior experiences influence teachers’ practices and current beliefs.
Theme Two: Focus on Issues of Fairness in Testing

Winke (2011) viewed the purpose of studying educators’ perspectives about assessment as a contribution to the consequential validity of an English language proficiency assessment. Winke (2011) agreed with Messick (1989) that tests should be more than just “statistically valid.” Messick (1994) and Linn, Baker and Dunbar (1991) argued for a broader sense of test validity, asserting that tests should also be fair, meaningful, cost-efficient, and developmentally appropriate. Thus, test validity includes the consequences of the test (Bachman, 1990; Messick, 1989). This focus on consequential validity provides an important reason for test developers and researchers to inquire into the perspective of teachers about tests.

Winke (2011) explored consequential validity using Michigan’s English Language Proficiency Assessment (ELPA), the state test that serves to assess English language proficiency in emerging bilingual students for accountability purposes. Her research questions inquired into educators’ opinions of the ELPA, and whether or not educators’ opinions were different depending on the specific demographic or teaching environment in which the ELPA was given. Winke (2011) used a three-part online survey with items that sought to investigate the social, ethical, and consequential dimensions of the validity of the ELPA test.

The central contribution of this study was to demonstrate the importance of broadening the concept of validity to include curricular and psychological effects of the test. Consistent with Messick, Winke (2011) urges us to extend the concept of validity beyond predictive validity and concurrent validity to include these social consequences. The perspectives of teachers are critical in this exploration, as they can easily point to these broader validity issues that are outside of the realm of test developers’ experience. Winke asserted that teachers’ perceptions should be indispensable to the development of a test’s validity argument. Since many testing companies
evaluate the validity of tests internally, testing companies would likely not encourage negative comments that practitioners may have regarding their tests. Thus, Winke supports the idea of external evaluations for a broader concept of test validity using teachers’ perceptions of the test. This would encourage test developers to be accountable not only to the institutions that require them, but also to the people the tests are meant to serve: students, teachers, parents, and the community.

*Theme Three: Focus on Policy Issues*

Harper et al. (2007) viewed the ‘equality of treatment’ perspective as problematic for emerging bilingual students as compared to an ‘equality of opportunity’ view, where social problems are considered systematic yet remediable. They argued that NCLB policy changes shifted the focus of equality in education towards equality of outcomes, such as test scores as indicators of equity. They posited that this focus on equality of outcome is different from previous policy initiatives in U.S. educational history that focused on equality of opportunity.

Harper et al. (2007) investigated the effects of NCLB policy on teachers of emerging bilingual students (specifically English as a Second Language, or ESL, teachers) by interviewing 52 teachers face to face or on the telephone. They conducted a cross-case analysis of the themes, and focused on one theme: teachers’ views on NCLB implementation with respect to reading. The authors found that, while teachers of emerging bilinguals expressed that NCLB had the positive effect of increased attention to the academic progress of emerging bilingual students, virtually every teacher had concerns about an excessive emphasis on reading in the curriculum, a limited concept of what constitutes reading, scripted instruction, and inappropriate assessment methods for emerging bilingual students. The focus on the standardized tests had reduced the
time allowed for instruction on topics other than reading. This led teachers to report that they felt like reading teachers instead of ESL teachers.

In all, the authors concluded that NCLB reforms have compromised rather than contributed to high-quality instruction for emerging bilingual students, particularly with respect to reading. They argue that this standardization has limited teachers’ capacity to decide what is best for students, and that teachers of emerging bilingual students can be forced to implement practices that are not helpful for these students, as is the case of reading instruction focused on phonics and oral fluency. They stress the importance of expert, autonomous teachers in helping children to develop literacy (Allington, 2002).

Jia et al. (2006) asserted from the beginning of their work that classroom assessments have more power to evaluate instruction and to identify students’ individual needs. They also stated that emerging bilingual students can benefit more from classroom assessment because it allows for assessment of language development to be integrated with content in a contextualized way. These authors appear to be uninterested in large-scale assessment, or at least its ability to drive quality instructional practices. Jia et al. (2006) instead focused on classroom assessments as a promising practice about which education professionals should learn more.

Jia et al. (2006) worked with 13 teachers of emerging bilingual students (namely, ESL teachers) in elementary and middle school classrooms across nine schools and four districts in Houston, Texas. The authors collected data through observations, interviews, and analysis of artifacts. They found that classroom-based reading assessments are central to teachers’ decisions about teaching and assessment of reading for emerging bilingual students, and that they inform the decisions teachers make about their students and the instruction best suited for them. However, they also found that teachers’ use of classroom-based reading assessments for
emerging bilingual students is mostly under the control of districts or school authorities and that statewide mandated standardized testing has taken precedence over teachers’ use of classroom-based reading assessment and distorted it. Finally, they found that teachers’ use of classroom-based reading assessments for emerging bilinguals is a “multi-faceted process,” working through the negotiation of many forces on teachers, including their own perceptions, external expectations from the district, the school, and the parents, and internal factors such as students, time, and materials.

The work by Jia et al. (2006) underscores the importance of studying the ways teachers think about classroom assessment with emerging bilingual students. They state that, “in order to have a comprehensive and profound understanding of classroom-based assessment, the voices of all the involved parties should be heard and addressed” (p. 426). They suggest that students, parents, and administrators should be interviewed to gain their insight on how assessment practices are used in their schools. The authors suggested that district and school administrators should focus on classroom and large-scale assessment equally, and that they should furnish teachers with more support with classroom reading assessments through an increased number of workshops or materials.

In a separate yet important work on the topic, Wright and Choi (2005) explained that the lack of Arizona teachers’ voices in the debate surrounding NCLB, Arizona LEARNS, and Proposition 203 is problematic, as teachers have first-hand knowledge and experience about how these policies are implemented in reality and how they affect emerging bilingual students’ education. Wright and Choi (2005) explained that teachers’ views are sometimes dismissed by policy makers since they are not necessarily representative of other teachers and schools throughout the state. Thus, the authors employed methods in order to accurately sample and
depict teachers’ perceptions of language and testing policies in the state of Arizona. Wright and Choi (2005) accomplished this task by conducting a telephone survey with a representative sample of experienced teachers of emerging bilingual students from the schools and districts with the largest populations of emerging bilingual students in the state. They chose to survey third grade classroom teachers, all of whom were working toward or already had an ESL or bilingual teaching endorsement, because they are dramatically affected by all language and testing policies: Arizona Proposition 203, Arizona LEARNS, and the federal NCLB Act of 2001. Their interview instrument contained selected response and open-ended interview questions.

Wright and Choi (2005) found that all of the teachers they surveyed believed that emerging bilingual students should learn English, but none of the surveyed teachers supported the idea that students should abandon their home language. They also found that, though all teachers expressed that emerging bilingual students and their teachers should be held accountable for student learning, almost all teachers were cautious about the equity of administering high-stakes content-area tests to students who are not yet proficient in the English language. The majority of teachers of emerging bilingual students also thought that high-stakes test scores are not very helpful in planning instruction for emerging bilingual students.

Furthermore, they found that 93% of teachers of emerging bilingual students reported that high stakes tests have not helped them focus on the linguistic and cultural needs of their students, and three fourths of teachers responded that high-stakes tests have not helped them to become a more effective teacher of emerging bilingual students. Unfortunately, 60% of teachers reported that they thought testing accommodations for emerging bilingual students were not allowed in their schools. The authors discussed how these changes led the majority of surveyed
teachers to claim that they have less satisfaction with their teaching career after the impact of
language, testing, and accountability policies.

To summarize, the literature about teachers’ views of assessment with emerging bilingual
students brings forth several important points. First, classroom assessment methods and tools are
instrumental in informing pedagogical practice for many teachers of emerging bilingual students;
it is important to support teachers in this endeavor (Jia et al., 2006). Second, teachers’
interpretations of assessment tools and policies are vital to developing valid and fair assessment
methods and policy practices for emerging bilingual students because of teachers’ direct
understanding of how the assessment tools, methods, and policies are carried out in schools
(Wright and Choi, 2005; Winke, 2011).

Finally, teachers are influenced by the community in which they teach, which can either
promote or detract from the facility with which they apply the theories they have learned about in
their teacher education classrooms (Rueda and Garcia, 1996). In fact, there are so many internal
and external factors that mediate teachers’ use of classroom assessment with emerging bilinguals
that it has been termed a multi-faceted process (Jia et al., 2006). This makes the factors that
influence teachers’ critical use of assessment difficult to identify. The multi-faceted nature of
classroom assessment use by teachers renders it both intriguing and challenging to study. In the
following section, I review what I believe to be two important influences involved in teachers’
critical use of assessment: critical thinking and critical pedagogy.
Section Two: Review of Critical Thinking and Critical Pedagogy

Critical Thinking

In order to better understand how critical thinking on the part of teachers may contribute to more valid assessment practices for emerging bilingual students, it is necessary to have a comprehensive view of what critical thinking entails. Though critical thinking has its roots in early philosophy through Socratic reasoning (Paul, Elder, & Bartell, 1997) and Buddhist thought (Herbert, 1993), the point at which critical thinking becomes essential to modern education in America is through the words of the fathers of democracy of the United States. In devising this social experiment called democracy, Thomas Jefferson wrote, “Above all things, I hope the education of the common people will be attended to… for the preservation of a due sense of liberty” (Glaser, 1941, p. 3). Glaser (1941) explained that the men who laid the foundation for this political structure were convinced that democracy would fail if the state did not cultivate a degree of social understanding and critical-mindedness that is necessary to judge public issues intelligently. From the birth of democracy in this country, critical thinking, though the name was not explicitly used, was already deemed as vital to the success of a community that worked together.

Thus, the concept of critical thinking has been highly significant since the founding of the institution of public education in this country. The reason its significance persists today is similar to the reason for its initial importance: to develop a degree of social understanding and critical-mindedness so that people might judge public issues intelligently within a democracy. However, the current environment of education has prioritized the concept of standardization instead, through programs like NCLB and Race to the Top. This priority of standardization can lead to classrooms with narrowed curriculum that focuses primarily on factual and procedural skills that
are easily measured by tests, instead of more conceptual knowledge like critical thinking (Abrams, 2004; Barksdale-Ladd & Thomas, 2000; Ravitch, 2010; Shepard, 2003; Strickland & Strickland, 1998; Taylor, Shepard, Kinner, & Rosenthal, 2003).

Indeed, we reward teachers and students who do well on standardized tests, which prioritize skills that are easily measured yet not necessarily those that are most highly valued. One of history’s most well-renowned intellects, Albert Einstein, agreed: “I believe in standardizing automobiles. I do not believe in standardizing human beings. Standardization is a great peril which threatens American culture” (Viereck, 1929). Perhaps the new Common Core standards and related assessment systems will motivate people to think more critically, since they are publicized as promoting higher order thinking skills. Regardless, we need to create a revival of teachers and students who are encouraged to engage in critical thinking within their communities. This movement would promote equitable assessment practices and improve instructional practices so future generations will continue to think critically and promote freedom through democracy.

Defining Critical Thinking

There seem to be as many definitions of critical thinking as there are authors who write on the topic. However, some of these authors fail to actually demonstrate the critical thinking skills that they implicitly are thought to possess. Some authors offer a definition that includes an exhaustive list of qualities that critical thinkers should possess instead of providing an overview that encompasses these qualities in a meaningful way (e.g. Facione, 2013; Foundation for Critical Thinking, 2013; Scriven & Paul, 1987). Some authors do not even attempt to explicitly define critical thinking in their volume on the topic (e.g. Bowell & Kemp, 2010).
In the same year as Scriven and Paul proposed their definition, Robert Ennis put forth a definition of critical thinking that was succinct and precise: “Critical thinking is reasonable, reflective thinking that is focused on deciding what to believe or do” (Ennis, 1987). This definition establishes a purpose for thinking, which is for people to decide which actions to take or what beliefs to hold, and it provides qualities to describe the kind of thinking that will help people to achieve this purpose. The reflective aspect of thinking that Ennis mentions was likely born out of Dewey’s (1933) work, which introduced the idea of reflection. Reflection refers to the idea that thinking cannot be impulsive, or controlled by instinct. Instead, one must slow down and reason about whether or not the instinctual reaction is indeed sensible. This is a concept in which Kahneman (2010) focuses on in his book, *Thinking, fast and slow*. Furthermore, Ennis (1987) delineated that there are two main aspects to critical thinking that must be taught (and we may infer all good critical thinkers should possess these two aspects): critical thinking dispositions and abilities.

In essence, to be critical thinkers, we must possess an inclination to think critically as well as to have the skills to think critically. This serves as an important distinction, since it demonstrates how having higher-order thinking skills or being an expert on Bloom’s taxonomy is not enough for one to be a good critical thinker; one must also use these skills regularly and integrate them into everyday decisions. Many other definitions take Ennis’ stated dimensions of critical thinking into account. In fact, authors before Ennis expressed these ideas as well. John Dewey, one of the most widely regarded authors in the field of critical thinking, expressed a similar claim about the need for positive “attitudes as well as skilled methods” (Dewey, 1933, p. 30). Glaser (1941) also used this concept when discussing the components that comprise critical thinking. Moon (2008) addresses the concept of academic assertiveness, which essentially
describes dispositions of critical thinking without using that terminology. Scriven and Paul (1987) call attention to the importance of having both critical thinking skills and habits based on intellectual commitment to regularly guide behavior. Hunter (2009) agrees with Ennis’ original dimensions of critical thinking as well.

In addition to Ennis’ definition of critical thinking, there are four other prominent authors who base their definitions of critical thinking in theory, rather than defining critical thinking without a supporting framework for the purpose of an educational text (R. H. Johnson, 1992). The definitions of critical thinking according to these authors, including Robert Ennis, are shown in Table 2, which was taken from R. H. Johnson (1992). There are similarities between the definitions; for example, reflective thinking is present across many of the definitions. Additionally, the skill of providing reasons to make good judgment cuts across the definitions. However, there are issues with every definition in Table 2, as R. H. Johnson (1992) describes, ranging from being too broad to capture thinking that is specifically critical in its definition (Siegel, 1988), to inviting negative connotations of the term ‘critical’ (McPeck, 1981).

Table 2

Definitions of Critical Thinking. Taken From R. H. Johnson (1992)

<table>
<thead>
<tr>
<th>Author</th>
<th>Definition of critical thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robert Ennis (1987)</td>
<td>“Reasonable, reflective thinking that is focused on deciding what to believe or do” (p. 10).</td>
</tr>
<tr>
<td>Richard Paul (1989)</td>
<td>“A disciplined, self-directed thinking which exemplifies the perfection of thinking appropriate to a particular mode or domain of thinking” (p. 214).</td>
</tr>
<tr>
<td>Matthew Lipman (1988)</td>
<td>“Skillful, responsible thinking that facilitates good judgment because it (1) relies upon criteria, (2) is self-correcting, and (3) is sensitive to context” (p. 39).</td>
</tr>
<tr>
<td>Harvey Siegel (1988)</td>
<td>“Thinking [that is] appropriately moved by reasons” (p. 32).</td>
</tr>
</tbody>
</table>
However, when reading these works on critical thinking and through critically-oriented
discussion, I found that there was a missing element to the definition of critical thinking as it was
described in the literature. The element of social context disappeared from the more current
works on critical thinking, though it was central in earlier contexts, such as in the era of Thomas
Jefferson’s presidency. The component of social context is vital to critical thinking in education
because in this field, teachers, students, parents, and administrators actively participate in a
community with a shared goal of helping students to learn. This social context is inextricable
from the thinking of individuals, since people constantly interact with each other in communities
through roles of apprenticeship, guided participation, and participatory appropriation (Rogoff,
1995). These interactions occur simultaneously to shape our learning and development as actors
within these communities (Lave & Wenger, 1991). Therefore, it is crucial to include the social
context in the discussion of critical thinking.

_Critical Thinking and Critical Pedagogy_

Critical pedagogy is chiefly concerned with developing “critical consciousness” of
systems of oppression, and bringing together reflection and action into praxis to promote
freedom from this oppression (Freire, 1970). This freedom from oppression necessitates that the
oppressed take part in their liberation; they cannot be passive in this process. Once the oppressed
become free, it is also important that they do not become oppressors of others, but instead
promote a society altogether free from oppression. In critical pedagogy, the term “critical” is
used in the same way critical theory employs it, though critical pedagogy and critical theory are
two distinct movements (Kaplan, 1994). Critical theory is a sociological perspective devised by
mostly German scholars in the early- to mid-1900s, whereas the critical pedagogy movement
identifies its’ beginnings in Paulo Freire’s work. Critical theory is a perspective that combines
Marxist, phenomenological, and psychoanalytical orientations. This theory critiques “lived social and political realities” in order to change those realities to ensure more freedom of thought and action (Kaplan, 1994, p. 207).

Critical pedagogy, then, makes use of the tools provided by critical theory in order to critique education, with the goal of ensuring that education leads to maximize human freedom (Kaplan, 1994). Thus, critical pedagogy works toward this goal through increasing the agency of students and teachers. This can be witnessed in Paulo Freire’s book, *Pedagogy of the Oppressed*, in which he discusses how education needs to support students to be active and autonomous actors, in opposition to the banking model of education (Freire, 1970; Kaplan, 1994). This is also observable in Giroux’s work entitled, *Teachers as Intellectuals*, where he urges teachers to view themselves with agency and power (Giroux, 1988).

Burbules and Berk (1999) demonstrate that both disciplines share common concerns about dispelling inaccuracies, distortions, and falsehoods. Critical thinking and critical pedagogy also share the belief that critical thought is not sufficient; one must pursue critical habits or action as a result of critical thought. For example, it is not enough to seek reasons and to try to be well informed (in the case of critical thinking) or to recognize injustice (in the case of critical pedagogy); one must also have a tendency to do such things (Ennis, 1987) or be moved to change the systems that create injustice (Marx, 1845).

With all of these similarities between the disciplines of critical thinking and critical pedagogy, however, the proponents of critical thinking and critical pedagogy use the term, “critical thought” within dissimilar contexts and with different premises for critical thought. Kaplan (1994) provides a helpful analysis of the different uses of “critical” as a terminology among scholars in the domain of critical thinking and scholars of critical pedagogy. She notes
that two nouns are related to the term “critical” as an adjective: “critique” and “criticism.”

Kaplan differentiates a criticism from a critique by explaining that a criticism entails the review of a work and its’ shortcomings in order to help the author improve the work, whereas a critique involves information on dimensions of meaning in the work about which the author of the work may be unaware. Kaplan (1994) then relates the meaning of this terminology to what critical thinking and critical pedagogy signify. She asserts that the critical thinking movement promotes the criticism of arguments, whereas the critical pedagogy movement advocates critique as a framework for critical thought about the world. Specifically, critical thinking generally provides critical feedback often within a framework of formal logic (e.g. Kaye, 2009), while critical pedagogy tends to provide this feedback by referring to issues outside of a pre-existing framework or system.

Burbules and Berk (1999) elaborate on more specific differences between critical thinking and critical pedagogy. They posit that the critical thinking tradition begins from the idea that individuals are free when they base their life choices on substantiated truth claims and are not subject to the undesirable control of unjustified beliefs, which is an issue that is nonpartisan in its’ nature (Burbules and Berk, 1999). The premise of critical pedagogy rests upon the concept of social justice, and people can become free when they realize that systems of belief and action have effects in power structures within society, and they pursue action to transform inequitable social relationships. The element of social context is abundantly clear in this tradition, to the point that authors from the critical thinking tradition assert that critical pedagogy crosses a line from teaching critical thought to teaching indoctrination about how society should be structured. Critical thinkers argue that teaching critical thought should allow for people to arrive at their own rational conclusions, whereas critical pedagogy nearly prejudices what the conclusions
should be (Burbules and Berk, 1999). Conversely, critical pedagogy views critical thinkers as already shrouded in indoctrination, and posits that people must be shown how these conditions have occurred in order to be able to think critically and justly about the world (Burbules and Berk, 1999). Thus, the tensions between the critical thinking and critical pedagogy traditions seem to be based on their situated political perspectives about whether or not Western thought is itself a system of oppression.

In the current work, the social context is of utmost importance when considering critical thought. While critical pedagogy begins from the premise that there are oppressive systems in the world, the theoretical framework of critical pedagogy involves developing critical thinking through pedagogical relations and collective action instead of having an emphasis on developing individual skills and dispositions, as is generally found in the tradition of critical thinking (Ennis, 1987; Glaser, 1941; Hunter, 2009; Kaye, 2009; Scriven & Paul, 1987). Critical pedagogy has social context embedded in the very fabric of the tradition, whereas critical thinking tends to include the social context as an addition that lies outside of the core of the discipline.

Critical pedagogy is vital to the current work not only because of its inclusion of social context in its literature, but also because critical pedagogy’s focus on social context leads to the explicit goal of equity in society. Critical thinking has no such goals; in fact, the field of critical thinking as a whole does not profess any objectives for justice, fairness, or equality in society. The mission of the tradition of critical thinking, according to many authors (Ennis, 1987; Lipman, 1988; McPeck, 1981; Paul, 1992; Siegel, 1988), is to develop, among individuals, the skills and dispositions necessary to become a critical thinker so that people may use good judgment in order to make logical and rational decisions. Logical decision making is clearly useful and important. However, critical thinking experts do not tend to relate logical decision
making to serve any larger purpose. Rather, logical decision making is an end in itself. For example, it is not an expressed priority to use logical decision making skills to work for a society that is more equitable or fair.

Conversely, critical pedagogy explicitly focuses on critical thought and action for the purpose of demanding equity and a balance of power among everyone in a society. Critical pedagogy’s goal of equity is essential for the current work. I acknowledge that the aim of encouraging teachers to use assessment critically for emerging bilingual students is to promote equity for all learners in the educational system. This study will only serve its purpose if it is able to address this aim of promoting equity in education.

Although the traditions of critical thinking and critical pedagogy do not tend to overlap, there are occasional exceptions (e.g. Brookfield, 1997; Giroux, 1994; Paul, 1992). Richard Paul (1992) is rooted in the tradition of critical thinking, yet sometimes bridges the gap between these two traditions. Richard Paul is regarded by Burbules and Berk (1999) as inclusive of social and contextual factors into the definition of critical thinking. Just as Freire (1970) promotes dialogue to characterize cultural action for freedom with the purpose of creating critical consciousness, Paul also explains that dialogical thinking is important to critical thinking practices. However, there are major differences between dialogue involving other people and dialogic thought, which takes place mostly within the individual. Though Paul (1992) explains that thinking from the perspective of other people is important to assessing claims of truth, he does not necessarily suggest that people solicit others’ actual points of view. Instead, Paul (1992) recommends that critical thinkers test frames of reference against each other dialectically, and “test the logical strength of one against the logical strength of the rest by appealing to standards not peculiar to any” (p. 270). Only if a person does not know how to make an argument from a certain frame of
reference should people resort to finding someone who can make the case for it. In this way, Paul
does not view dialogue between people as necessary to develop critical thinking from a frame of
reference with which one is familiar.

Though Paul (1992) does not inquire directly into other people’s points of view, he is
earnestly concerned with issues of sociocentrism. He argues that if people do not consider other
standards of evidence, then they could falsely reject other credible points of view that may be
otherwise silenced (Burbules and Berk, 1999). Sociocentrism, according to Paul (1992), leads to
errors in thought. Paul believes that critical thinking has all of the tools to overcome
sociocentrism; however, he does not see a need to shift the focus of critical thinking to a situated
perspective that reflects the nature of human thought, as described by sociocultural thinkers
(Lave & Wenger, 1991; Rogoff, 1995; Vygotsky, 1978; Wenger, 1998). Paul is more socially
minded than other critical thinking authors, yet he still assesses critical thinking skills in the
context of individuals who have a “strong sense,” “weak sense,” or no sense at all of these
values.

Giroux (1994) also bridges the gap between critical pedagogy and critical thinking. He
comes from the critical pedagogy perspective, yet he also writes about the practice of critical
thinking for students. Giroux (1994) rejects what he labels the “internal consistency” position,
which refers to the definition of critical thinking as the analysis and development of reading and
writing from the viewpoint of formal, logical patterns of consistency. He argues that, even if an
argument is consistent within the formal system of logic taught to students, people need to look
outside the confines of what has been presented to see if there are issues that have been excluded.
He believes that people need to challenge the assumptions that lie within the system, for
instance, by asserting that knowledge is inseparable from human interests and values and by proposing that facts and theory are related to each other.

Thus, Giroux (1994) redefines critical thinking in a way that does not promote the teacher as the carrier of knowledge and logical argumentation necessary to think critically but, instead, as a facilitator who mediates between students and the larger social reality. He asserts that teachers should challenge students to think about which types of knowledge are legitimate, in addition to evaluating the knowledge that comes from information already provided to students. This approach does not interpret students’ conclusions for them. Instead, it opens the door for students to be critical not only about the information students are already given to analyze, but also with regard to information that is omitted (Giroux, 1994).

Critical Thinking and Second Language Acquisition

Uniquely positioned as an expert in the field of second language acquisition, Cummins (1989) also wrote from the perspective of critical pedagogy to disrupt power structures in educational assessment practices. Cummins (1989) believes that, since people learn through social interactions, teachers can use their own growth and learning in the area of critical thinking to influence their students’ development in the area of critical thinking as well. Cummins (1989) writes that when teachers begin to “critically analyze the forces that disempower them within the classroom, they automatically take the first steps toward the empowerment of their students” (p. 36). Therefore, Cummins professes that positive change occurs when teachers actively challenge the structure within which they function. He directly relates this to the field of assessment by asserting that, when there are more empowered ESL students and teachers, it becomes possible in the long term to liberate the educational system from the “straightjacket of standardized tests” that currently operate as tools to oppress teachers and suppress learning (p. 35). The metaphor
TEACHERS AS CRITICAL ASSESSMENT USERS FOR EMERGING BILINGUALS

highlights the idea that standardized testing in its current form may limit teachers’ abilities to emphasize certain areas of cognition such as critical thinking, creative writing, and other endeavors that do not lend themselves as easily to multiple-choice and short answer item formats.

Cummins (1989) argues that, if nothing else, it is important for students and teachers not to be passively controlled by higher levels of educational hierarchy. He believes an active stance would allow teachers to demand assessment systems that include all the curriculum objectives present in a sociocultural theory-based pedagogy. He suggests a more feasible goal—that if there are no tests available on the market to assess objectives such as critical thinking, then teachers should pressure the school district to develop these assessments themselves. He also states that these assessments should be given equal or greater weight than standardized tests in the overall assessment system.

Granted, even Cummins’ “more feasible” goal of asking teachers to pressure school districts to create assessment systems that assess objectives such as critical thinking or creative writing would take considerable grassroots efforts. Standardized testing systems have skyrocketed in terms of their overall importance and weight with respect to the evaluation of students, teachers, and schools themselves. The heightened value of assessment in the current educational environment in the U.S. only serves to increase the importance of the goal to implement assessments that reflect the understandings from second language acquisition theory and that assess all students equitably. This notion is consistent with Giroux’s (1994) belief that students will learn little about critical thinking if they are exposed to a language, belief, and value system whose message conveys that they are disadvantaged.
The Assessment of Critical Thinking and Critical Action

How, then, would one determine if someone is thinking critically or not, from the perspective that critical thinking is not an intrinsic, heritable trait, but a trait that develops with experience through social interaction (Martinez, 2000)? In an article called “Assessing critical thinking,” Brookfield (1997) expands on how one might develop assessments to measure critical thinking skills, given that critical thinking is a skill cultivated through social interaction. Unlike Cummins, Brookfield believes that critical thinking is a skill that can only be effectively developed within adult minds. Regardless of the age of the students developing critical thinking skills, I believe that his suggestions for the approach to assessment of these critical thinking skills would apply.

Brookfield (1997) posits that critical thinking serves to help people examine assumptions that undergird people’s thoughts and actions, such as how we interpret relationships of power in our lives, and how hegemonic assumptions affect us. In order to determine if people are actually engaging in the work of examining their assumptions within a certain context, Brookfield believes that it is necessary to use assessments that include social interaction or reflection on interaction. He views this as necessary because he claims that critical thinking is undoubtedly a social process. Thus, Brookfield believes that there is no single assessment that can be applied to any situation to assess critical thinking skills, since critical thought is embedded in one’s social, linguistic, and cultural surroundings. Assessments must be specific to their context, and this contextualization is best accomplished by peers who can operate as critical mirrors, reflecting and highlighting the assumptions that might not have been apparent to a person by themselves or to their instructor as an individual. Additionally, assessment of critical thinking needs to allow
people learning about critical thinking to document, demonstrate, and justify their engagement in critical thinking (Brookfield, 1997).

Due to his belief that critical thinking is context-specific and embedded in one’s social, linguistic, and cultural environment, Brookfield ultimately suggests that there is no single appropriate way to assess critical thinking skills. He does provide options for instructors to consider when they plan to assess the critical thinking skills of their students. One of the options includes a weekly recording of critical incidents that take place in students’ daily lives, coupled with questions to reveal the assumptions that students make in order to interpret the events. Another possibility is to engage students in a critical debate about a controversial issue, and then to ask students to respond to questions about the events that transpired during the debate.

Yet another possibility is to engage a student in the role of a storyteller in depicting an event that happened to them, while other students help the storyteller to examine their assumptions. In this scenario, a different student takes the role of an umpire in order to point out when people pass judgment on the storyteller instead of asking questions and giving descriptions of the account they heard. In all of these suggestions for assessment of critical thinking skills, there is a personalized, contextualized basis for the examination. Although Brookfield cautions instructors against following these ideas for assessment exactly, these suggestions provide real, practical examples of how assessment of critical thinking skills within a social context can be accomplished.

*Reflection on Critical Thinking and Critical Pedagogy*

To reflect on all of the literature on critical thinking and critical pedagogy, in this section I distill the many attributes that comprise critical skills and habits of mind into the core components that I employ in this study. A common thread that is woven throughout various
authors’ conceptions of critical thinking and critical pedagogy is the necessity for people to develop both the skills and the disposition to think critically (Dewey, 1933; Ennis, 1987; Glaser, 1941; Moon, 2008), or by engaging in both reflection and action to develop praxis (Freire, 1970; Marx, 1845). Another key component of critical thinking and critical pedagogy includes the act of questioning: either by demanding reasons (Siegel, 1988) or challenging dominant rhetoric (Apple, 2008; Giroux, 1988). People may question their own assumptions to critically examine and reflect upon their own actions (Brookfield, 1997), or they may question the legitimacy of information and knowledge provided to them by others (Giroux, 1988, 1994).

Questioning is central to critical thinking and critical pedagogy. Questioning can be viewed as both a skill which one can develop, and it can also be seen as a habit of mind that one may strive continually to perform. Additionally, reflection is a habit of mind that provides the time and space in which people may question their practices and the practices of the world around them (Brookfield, 1995; Dewey, 1933). In order to think critically, it is important to base one’s judgment on reasons that are sensitive to the context of the situation (Lipman, 1988). If one provides logical reasons for making a decision, the decision is more easily substantiated. Also, if the reasons are sensitive to context, one’s decision making process will more likely account for differences across situations, and then the critical thinker can correct their criteria used in making decisions accordingly.

Finally, the goal of using critical thinking to promote equity for all students in education is an important critical attribute in this work. While the objective of equity in education is only expressed in the literature of critical pedagogy, it is a mission that is essential to the current study and thus it is incorporated into the conceptual framework that is used to help identify when and how an educator uses assessment critically.
Specifically in the context of assessment with emerging bilingual students, there are several pertinent critical skills and habits of mind. A profile of teachers who use assessment critically is necessary in order to then assess if teachers are demonstrating these characteristics. Thus, teachers who are critical users of assessment should have the habits and skills listed below:

*Questioning*

1. *Question* their own assessment practices to determine if these practices serve their intended purposes well for all students, particularly emerging bilinguals.
2. *Question* the assessment practices that are required and recommended by the school, district, state, and nation in order to determine if they serve their intended purposes well for all students, particularly emerging bilinguals.
3. *Engage* in life-long learning to improve one’s teaching and assessment practices for all students, particularly emerging bilinguals.

*Decision making based on substantiated reasons*

4. *Understand* the criteria that contribute to appropriate, high-quality assessment tools and methods, according to research literature. Specifically, a teacher understands that an appropriate test depends on the relevance and significance of the results for the teacher and their students, the value of the assessment in promoting subsequent instruction, the validity and reliability of the results of the test, the fairness of the test to all students, and the efficiency of the test administration, for example.
5. *Select or develop* appropriate assessment tools and methods in order to obtain a valid and reliable picture of all students’ skills, particularly emerging bilinguals, in the relevant content area for the class.
(6) Evaluate appropriate assessment tools and methods in order to obtain a valid and reliable picture of all students’ skills, particularly emerging bilinguals, in the relevant content area for the class.

(7) Administer, score and interpret assessment tools in order to obtain a valid and reliable picture of all students’ skills, particularly emerging bilinguals, in the relevant content area for the class.

(8) Use assessment methods and tools to promote student learning.

(9) Employ assessment methods and tools to inform one’s instructional practice.

(10) Use assessment results to make substantiated decisions about emerging bilingual students, teaching, curriculum, and school improvement for all students, particularly emerging bilingual students.

(11) Use assessment tools to develop valid grading procedures for all students, particularly emerging bilingual students.

Reflection

(12) Reflect on the assessment tools and methods they [the teachers] use in their classrooms, after each use, to determine if they are serving the purpose of identifying what all students know validly and reliably.

(13) Consider the advantages and limitations of various assessment tools used with linguistically and culturally diverse populations in order to assess all students equitably.

(14) Help students learn to reflect upon and monitor their own learning with assessment tools.
Decision making that is sensitive to various contexts

(15) Consider how different students, disciplines, modes of language, and contexts can all change an assessment tool’s ability to identify what students know validly.

(16) Employ a variety of assessment methods and tools to gain insight about students in different contexts in order to make more informed decisions about all students, particularly emerging bilinguals.

Social context

(17) Communicate assessment results to emerging bilingual students, their parents, and other educators.

(18) Participate in larger forums on assessment with emerging bilingual students.

(19) Share knowledge and understandings about assessment with emerging bilingual students with colleagues.

(20) Discuss, as a community of teachers, which assessment tools and methods are most appropriate for the current students.

Goal of using critical skills to promote equity in assessment for all students

(21) Recognize unethical, illegal, or inappropriate assessment methods and uses of assessment information, both in large scale and classroom assessment.

(22) Take action when unethical, illegal, or inappropriate assessment methods and information are used.

(23) Support and collaborate with families of students to ensure their voice is heard with regard to issues of assessment for students, particularly emerging bilinguals.

(24) Help students to become more critical in their daily lives.
Section Three: History of the Standards for Teachers’ Skills in Assessment and the Teaching of Emerging Bilingual Students

Assessment Standards

A literature review of the established assessment skills in which teachers should be proficient is vital to the study of teachers’ use of assessment, critical or otherwise. Assessment is a fundamental part of instruction in the classroom that can help teachers to improve their pedagogy and that can help students take ownership of their learning as well. Additionally, the past decade has witnessed a national movement to focus on assessment for accountability reasons through the implementation of NCLB and through the financial incentives provided by the Race to the Top program. This focus on assessment provides incentive to increase teacher competency with regard to assessment, to be able to both use assessment as a tool to improve instruction and to interpret assessment systems when they are implemented on a large scale.

Unfortunately, there have not been many updated standards or formal guidelines developed for the purpose of establishing the assessment skills that teachers should possess, despite their growing importance in our current educational environment. The most prominent assessment skills standards designed for teachers from grades K-12 were developed in 1990. In their chapter on “The Development of Standards for Teacher Competence in Educational Assessment of Students,” Sanders and Vogel (1993) posit that teachers should be skilled in assessment in seven main areas. These standards were developed by measurement professionals, through the combined efforts of the National Council on Measurement in Education (NCME), the American Association of Colleges of Teacher Education (AACTE), the American Federation of Teachers (AFT), and the National Education Association (NEA) to address the areas in which it was vital for teachers to be competent with regard to their assessment practices.
These standards state that teachers should be able to: (1) choose appropriate assessment methods for instructional decisions, (2) develop their own appropriate assessments for instructional decisions, (3) administer, score and interpret their own or externally-developed assessments, (4) use assessment results to make decisions about students, teaching, curriculum, and school improvement, (5) use assessments to develop valid grading procedures, to participate in larger forums on assessment, (6) communicate assessment results to students, parents, and other educators, and (7) recognize unethical, illegal, or inappropriate assessment methods and uses of assessment information.

Considering that these standards were written in 1990, they pay relatively more attention to linguistic and cultural diversity than one might have expected. The first standard suggests that teachers should consider “cultural, social, economic, and language backgrounds of students” (Sanders and Vogel, 1993, p. 56) in choosing appropriate assessments for instructional decisions. In Standard 6, the authors stress that the interpretation of student assessments should be “moderated by the students’ socioeconomic, cultural, language, and other background factors” (p. 59) and that teachers should be able to explain that assessment results do not imply that a student’s background limits their educational development in the long run. Standard 7 also capitalizes on the need for fairness in assessment, and encourages teachers to attempt to have inappropriate assessment practices discontinued whenever they are discovered.

Though these standards do not directly address critical thinking, it is apparent that the authors indirectly value critical thinking, because they explain how a teacher who is a critical user of assessment would act. For example, Standard 7 explains that teachers will (a) be well versed in their ethical and legal assessment responsibilities, and (b) they should take action to discontinue assessments when they conflict with their concept of fairness. This shows that
Teachers are expected to both have critical skills and understanding, while also being expected to have the disposition to take action when they see that assessments are not being used fairly. This seventh standard, in particular, touches on both critical thinking and critical pedagogy. There could be improvements on these standards; however, by focusing more intensively on issues of linguistic diversity and critical thinking.

There are other published assessment standards. However, they all have shortcomings with regard to effectively reaching the audience of P-12 teachers in the current educational context. The Assessment Skills and Knowledge Standards, developed by the American College Personnel Association (ACPA), were targeted at professors and departments at the college level (Mitchell, 2006). The American Psychological Association (APA), National Council of Measurement in Education (NCME), and American Educational Research Association (AERA) all worked together to develop Standards for Educational and Psychological Measurement in 1999 and updated them in 2014 as well. However, these standards targeted test developers and educational researchers rather than teachers. The Code of Professional Conduct in Measurement, which was developed in 1995 by the NCME, targets teachers as one segment of its audience. In contrast, the Code of Professional Conduct also targets test developers and companies. In 2002, the Student Evaluation Standards, created by the Joint Committee on Standards for Educational Evaluation (JCSEE), include 28 standards to help guide P-12 teachers in their evaluation practices (JCSEE, 2003). These standards target teachers as their primary audience, and the term “evaluation” is used similarly to “assessment.” However, these standards omit any discussion of large-scale assessment or teachers’ role within the process of large-scale assessment systems.

With the exception of the Student Evaluation Standards, none of the standards with P-12 teachers as an audience have been updated since the turn of the 21st century, despite the rise in
the importance of assessment methodology in the educational system. The APA/NCME/AERA standards also stand out as an exception since they were updated this past year, but there is another issue in addition to the fact that these standards are mostly outdated. These standards are not affordable or easy to access. For example, the APA/NCME/AERA 2014 standards are 70 dollars to purchase, and the Student Evaluation Standards are priced at 40 dollars. Unless schools purchase these standards for their teachers, they will likely not reach the majority of teachers at the P-12 level.

Another concern regarding teacher standards for assessment methodology is the issue of usability. The 2002 Student Evaluation Standards are a quality resource, yet the 28 standards comprise an entire volume with 208 pages of descriptions and examples. This level of detail may be excessive if usability is a primary concern. The seven standards laid out in the 1990 Standards for Teacher Competence in Educational Assessment of Students have a high level of usability, yet are outdated.

Brookhart (2011) suggested that the 1990 assessment standards (Sanders & Vogel, 1993) are the optimal grain size to guide teachers in their assessment practices, since other guidelines that have been developed since the 1990 standards are either overly broad (e.g. McMillan, 2000; National Board for Professional Teaching Standards, 2001; National Council for the Accreditation of Teacher Education, 2008) or too detailed (e.g. Joint Committee on Standards for Educational Evaluation, 2003). Thus, Brookhart (2011) believes that the 1990 standards fulfilled their purpose; however, they currently need to be updated to reflect certain changes in the focus of recent educational policy and research. She posits that new standards should account for the advances that have taken place with respect to formative assessment as well as the role that assessment plays in the current environment of “standards-based reform” and accountability.
Brookhart (2011) even suggests potential standards that include these new concepts. However, due to the fact that her article is not freely accessible to a P-12 educator audience, it is unlikely that her proposed standards will reach most teachers as well.

Thus, well-crafted assessment standards for P-12 teachers exist, yet they are either outdated or they are not readily accessible and affordable for many teachers. The lack of assessment standards with these qualities demonstrates that, while there has been an increased national focus on assessment for the purposes of accountability and standards-based reform, there has not been a corresponding increase in focus on preparing the nation’s teachers to succeed in the context of these new assessment systems.

*Standards for Teachers of Emerging Bilingual Students*

If teachers are to be critical users of assessment with emerging bilingual students, it is important for teachers to have skills in helping students to acquire a new language. There are two recent and prevalent sets of standards that serve to inform teachers about the skills in which they should be proficient in order to support their students in simultaneous content and language learning. The Teaching of English to Speakers of Other Languages (TESOL) International Association worked with the National Council for the Accreditation of Teacher Education (NCATE) to create standards for the certification of initial TESOL programs in P-12 ESL Teacher Education. The TESOL standards were first employed in 2001, then updated in 2009 (Staehr Fenner & Segota, 2012). The TESOL standards are important, as they are used by over 200 institutions of higher education as the framework for their English as a Second Language (ESL) programs (Staehr Fenner & Segota, 2012).

The National Board for Professional Teaching Standards (NBPTS) also has a set of standards for the pedagogy of English as a New Language (ENL) that were first developed in the
1990s and were updated in 2010 (Staehr Fenner & Segota, 2012). Yet, these standards were developed for individual teachers instead of teacher education programs. Also, the audience for the NBPTS ENL standards is the accomplished ESL teacher with at least three years of experience in the classroom. Teachers may voluntarily work towards certification under the NBPTS standards, whereas the TESOL standards tend to be used as a basis for teacher education programs for mandatory initial licensing purposes (Staehr Fenner & Segota, 2012).

Other similar standards available include the Center for Research on Education, Diversity & Excellence (CREDE) Standards for Effective Teaching Practice and the National Association of Bilingual Education (NABE) Professional Standards for the Preparation of Bilingual / Multicultural Teachers (Menken, Antunez, Dilworth, & Yasin, 2001). However, these standards were published in 1998 and 1992, respectively, so they do not include information that is as relevant to the current context of education for emerging bilingual students. Additionally, the CREDE standards do not focus on linguistic issues; rather, their intent is to highlight effective and equitable pedagogical practice for all students, whether they belong to a majority or a minority population. Therefore, I review the TESOL standards and the NBPTS ENL standards in more depth to determine the specific areas in which teachers of emerging bilingual students are expected to be proficient.

The TESOL Standards state that teachers should know, understand, and use the major theories, concepts, issues, and research related to (1) the structure and acquisition of language to help ELLs develop language and literacy and achieve in the content areas, (2) the nature and role of culture and cultural groups to construct supportive learning environments for ELLs, (3) planning, implementing, and managing standards-based ESL and content-area instruction, and (4) assessment using standards-based procedures with ELLs (TESOL/NCATE, 2010). The
standards also state that teachers should (5) keep up to date with new instructional methods, research, advances in the field of ESL, and issues surrounding educational policy as well as to demonstrate an understanding of the history of ESL teaching (TESOL/NCATE, 2010). Teachers should “use this information to reflect on and improve their instruction and assessment practices,” and they should “work collaboratively with school staff and the community to improve the learning environment, to provide support, and to advocate for ELLs and their families” (TESOL/NCATE, 2010, p. 68).

TESOL’s Standard 4 is overtly applicable to the current study, as it details the understandings that teachers should possess with regard to assessment methods for emerging bilingual students. The standards delineate that teachers should understand the advantages and limitations of assessment methods and testing accommodations for emerging bilingual students, as well as the key indicators of quality assessment techniques. With regard to language proficiency assessment, the standards explain that teachers should be able to demonstrate an understanding of how to use assessment to track language development on the part of students, in addition to how to inform their subsequent instruction for emerging bilingual students. They stress that emerging bilingual students’ language proficiency level should be measured using multiple sources of information, including performance-based measures in order to assess students’ communicative competence in various contexts across the curriculum. The standards also explicitly state that assessment should be used to promote as well as to measure learning (TESOL/NCATE, 2010).

The supporting explanation of Standard 4, Section C of the TESOL standards cautions teachers that some of the classroom assessments which are designed for native speakers, such as independent oral reading assessments, may be “uninformative or misleading” when they are used
with emerging bilingual students (TESOL/NCATE, 2010, p. 64). This statement is significant to the current work, in that the standards implicitly advise teachers to use good judgment when selecting assessment tools and when making inferences about emerging bilingual students from assessment. Finally, Standard 4 stipulates that teachers who exceed the standard should share their understandings about assessment methodology with other colleagues. This communicates that it is not only important to engage in critical reflection by using good judgment in one’s own assessment methodology, but it is also important to advocate that others use good judgment as well. This standard describes the critical action piece that is an important component in critical use of assessment.

TESOL Standard 5 is also relevant to this work, since it continues the demand for teachers to act as advocates for their emerging bilingual students and their families. Specifically, the supporting explanation in Standard 5 states that teachers need to “advocate for appropriate instruction and assessment by sharing their knowledge of ELLs with their general education and content-area colleagues and the community” (TESOL/NCATE, 2010, p. 71). This notion of advocacy is a vital component of critical use of assessment by teachers, since it challenges teachers to follow through on their reflections about issues surrounding new assessment methods and educational policy.

The NBPTS Standards for English as a New Language discuss critical thinking and reflection at length in the 2010 edition. Central to NBPTS’ framework are five core propositions of all accomplished teachers. These propositions span the standards for teachers of all subject areas and grade levels, not only English as a New Language. These central tenets state that: (1) accomplished teachers need to be committed to their students and their learning, (2) they know the subjects they teach and how to teach those subjects, (3) they are responsible for managing
and monitoring student learning, (4) they think systematically about their practice and learn from their experience, and (5) they are members of learning communities. The fourth core proposition of the NBPTS standards specifies that the accomplished teacher thinks systematically about their practice, which involves critical reflection that is particularly important to the current study. In fact, the description of the core proposition explicitly suggests that, “striving to strengthen their teaching, accomplished teachers examine their practice critically” (NBPTS, 2010, p. 12). This indicates that NBPTS views a critical examination of one’s practice as vital to every teacher’s practice, since this proposition is included at the beginning of every set of standards across subject areas and grade levels. The description of this proposition specifies that teachers must read about current issues, question them, and incorporate new findings into their practice.

NBPTS standards specifically for English as a New Language include the demand for knowledge of students; understanding of various cultures and diversity; knowledge of the English language and of English language acquisition; the ability to make home, school, and community connections; understanding of instructional practice and of assessment; and, finally, being a learner and a professional leader and advocate. Standard 7 describes how accomplished teachers should approach assessment methodology: “Accomplished teachers of English language learners employ a variety of practices to assess their students appropriately. They use assessment results to shape instruction, to monitor student learning, to assist students in reflecting on their own progress, and to report student progress” (NBPTS, 2010, p. 77).

The description of the first statement in Standard 7 asks that teachers recognize the advantages and limitations of a range of assessment methods, and that emerging bilingual students often have skills that do not emerge in certain types of assessment processes. Standard 7 also demands that teachers address the possibility that some assessments may favor the dominant
culture. For these reasons, the standards suggest that accomplished teachers should not rely on any single method of student achievement. In order to fulfill these demands, teachers must employ critical thinking skills to evaluate the strengths and weaknesses of assessment tools and how assessment tools may work together in order to provide a full, more robust picture of a student’s academic progress.

To summarize, teachers’ critical thinking skills and activism are emphasized in the 2010 editions of the TESOL and the NBPTS standards. The critical use of assessment with emerging bilingual students is a clear priority for the professional organizations that develop these standards and the universities who employ these standards as a foundation for their programs. The standards are still quite vague; however, so teachers might need more specific information that they could use in order to develop these skills fully. Just as other skills are developed, critical thinking is learned and acquired over time. Once teachers’ critical thinking skills about equitable assessment methods are developed, they might need fewer guidelines regarding equitable assessment practices. When teachers begin this endeavor; however, it is helpful to have guidelines. There are few formal guidelines specifically created for teachers with respect to the appropriate assessment of emerging bilingual students (e.g. Gottlieb, 2006; Valdez Pierce & O’Malley, 1996; Brantley, 2006). While assessing students equitably is a priority in the standards for teachers of emerging bilingual students, there are not many resources that are readily available to assist teachers in this venture.
Chapter 4
Methodology

Strategy of Inquiry

This is a mixed-methods study, which employed mostly qualitative methods to gather information from a small number of teachers through interviews and observations. This study also incorporated quantitative methods by using a survey with a sufficiently large number of teachers that ensured statistical power. The survey was distributed among teachers of emerging bilingual students to identify the extent to which a broad range of teachers think and act critically with regard to assessment of emerging bilingual students. Statistical analyses were performed to identify which aspects of critical thinking and critical action, in regard to assessment of emerging bilingual students, are regularly implemented by teachers and which present challenges.

In-depth qualitative interviews with teachers allowed for an analysis of the processes that teachers employ in order to reflect on and use assessment critically. Interviews also shed light on the factors that affect teachers’ critical use of assessment. A set of qualitative codes were developed and used to analyze the data from teachers’ responses. The categories were created based on the conceptual framework for the current study. Additionally, these categories were employed to identify the location of each teacher on rubrics intended to evaluate the extent to which a teacher thinks critically about the assessment of emerging bilingual students, and the extent to which a teacher uses assessment critically for these students. Finally, to triangulate data, three observations of each of the six focus teachers’ classrooms occurred purposefully during class periods when the teacher indicated that classroom assessment would occur, in order to observe how teachers actually use assessment in their own classrooms.
Participants and Design

Table 3 shows the research design. I targeted 82 current teachers in the state of Colorado who are practicing in public or charter schools with at least a few emerging bilingual students in their classes. I selected practitioners teaching strictly in grades 3-8 for the interviews, as these are the grades in which large-scale testing may have a more significant impact on classrooms. Also, a greater proportion of students in these grades are either transitioning to instruction in English or actively learning English as a new language at this point in their educational experience. I selected practitioners teaching in Grades K-12 for the survey to obtain a wider perspective across various classrooms on the issues presented in my research questions.

All of the teacher participants had emerging bilingual students in their classrooms. However, only half of the interview participants (5) attended an educational program at the college level in order to obtain a “Culturally & Linguistically Diverse Education” or “Culturally & Linguistically Diverse - Bilingual Education Specialist” (CLD) teaching endorsement. The other half of the interview participants (5) did not attend an educational program at the college level in order to obtain a CLD teaching endorsement, or they did not have a CLD teaching endorsement. There were two interviewed teachers who did obtain a CLD teaching endorsement simply from taking the Program for Licensing Assessments for Colorado Educators (PLACE) test, which was not the type of experience that could change someone’s perspective in the same way that an educational degree program would. These two teachers were then counted as teachers without CLD program experience.

The criteria for the survey participants were identical to the criteria for the interview participants, aside from the fact that the grade level criteria was extended to K-12. Half of the survey participants (30) attended or are attending an educational program at the college level in
order to obtain a CLD teaching endorsement. The other half of the surveyed teachers (30) have not attended an educational program at the college level in order to obtain a CLD teaching endorsement. I purposely structured the survey participant selection with these two distinct categories of teachers. These categories were chosen in order to obtain participants for the survey who are likely to have varying skill levels and experience with regard to critical use of assessment for emerging bilingual students.

There were 12 survey participants who obtained a CLD teaching endorsement without naming an educational program they attended in order to obtain the endorsement. These 12 responses were discarded from the survey dataset due to the fact that the background of these teachers did not fit the criteria for CLD status in the study. These 12 participants could have potentially been included in the group of teachers with non-CLD status, though it was impossible to tell if they simply forgot to name the educational program they attended to obtain a CLD endorsement, or if they truly did not attend an educational program to obtain their CLD endorsement. Due to this uncertainty, their data was discarded. There was also likely to be more contrast in the survey results by only including two strict groups of teachers: (1) those who had received a CLD endorsement through experience in an educational program and (2) those who had not received a CLD endorsement at all. There were equal proportions of surveyed teachers disaggregated by CLD status (n=30 in each group) without including these 12 participants.

Within each group of teachers designated by CLD status, there are three different groups of participants: one group of participants (n=30) took a survey only. A smaller, separate group of participants (n=5) participated in two interview sessions, and a final sub-group of the interview participants (n=3) also participated in three classroom observations. This final sub-group of three CLD-experienced and three non-CLD experienced teachers is termed “focus teachers” in this
work. Thus, there were 35 teachers total from each group of teachers designated by CLD status, leading to a total of 70 participants in the study after the data from the 12 teachers with uncertain levels of CLD experience were discarded.

In order to find teachers who fit the selection criteria, I surveyed many current and former Master’s students from the Educational Equity and Cultural Diversity (EECD) program at the University of Colorado at Boulder. The Master’s EECD program focuses on education for emerging bilingual students and for students with special needs, and the emphasis on linguistically diverse education within the program qualifies one for the CLD endorsement in the state of Colorado (The Colorado Department of Education, 2013; The School of Education at the University of Colorado at Boulder, 2013). I exclusively consider Master’s students who have an emphasis on linguistically diverse education when I further discuss or refer to Master’s EECD students in this document. I also surveyed teachers whose program experience took place at other institutions. I recruited these other survey participants by asking teachers to take my survey at CLD-related events, such as the Colorado Association for Bilingual Education (CABE) conference, and by asking contacts to spread the news of the survey to other teachers in Colorado.

An assumption underlying this selection process is that, due to teachers’ exposure to issues of language and assessment during their time at the University, these teachers are more likely to have developed higher skill levels of critical use of assessment for emerging bilingual students than practitioners who have not been in the Master’s EECD program. I considered teachers who were currently enrolled in their CLD educational program experience as well as those who had completed the program as part of the CLD-experienced group, since the
experience of the CLD program is fresh on participants’ minds when they are currently enrolled in the program.

To recruit participants for the second large group of teachers who do not have or expect to have a CLD endorsement, I contacted teachers who are former students of the undergraduate teacher education program at the University of Colorado at Boulder who do not have or expect to obtain a Master’s in EECD or its equivalent at another University. I also attended conferences that current practitioners of certain content specialties, such as science, were required to attend. Since these non-CLD-experienced teachers are less likely to have the same level of formal exposure to issues of language and culture in assessment as those teachers who participated in a CLD educational program, it is reasonable to predict that their skill level as a critical user of assessment with emerging bilingual students is not likely to be as high as those teachers with CLD program experience. Table 3 also shows the CLD program experience status of the study participants.

Table 3

Research Design

<table>
<thead>
<tr>
<th>Participant group</th>
<th>CLD educational program experience</th>
<th>No CLD educational program experience</th>
</tr>
</thead>
</table>
| Survey participants
a                                 | 30                                 | 30                                    |
| Interview and classroom observation participants (focus teachers) | 3                                  | 3                                     |
| Interview participants only                           | 2                                  | 2                                     |
| Total (n=70)                                          | 35                                 | 35                                    |

*a An additional 12 participants were surveyed, but were not included in the data because they did not meet the set criteria.

I engaged in the interview process with participants prior to the survey administration. This decision was made based on the idea that the information gathered from interview participants could inform the development of the survey items. I used the experience from
conducting interviews to refine the survey items, based on an informed understanding of the items that caused interview participants to share ideas relevant to the study’s research questions. Though I had initially planned to use the survey results to find interview participants with a range of skills in the area of critical thinking, I concluded that the survey would reach the highest number of participants and that it would be more important for the survey instrument to be informed by the interview process.

**Researcher Role**

I administered the aforementioned survey to 72 teachers, interviewed ten additional teachers, and observed six of those teachers’ classrooms. Therefore, I would be remiss to omit a description of the role I played in this study. In order to gather the kind of information necessary to answer my research questions, I needed to develop a professional relationship with the teachers I interviewed. Thus, all of my interactions with the teachers I interviewed influence the way in which I interpret their level of critical thinking and critical action holistically.

Additionally, the motivation for this study derives from my own experience as a teacher of emerging bilingual students. Since I enrolled in this doctoral program with the goal of improving assessment practices for emerging bilingual students such that they are more equitable, I am invested in finding ways to achieve this goal. This specific topic of study originated from my own perceived successes and failures in thinking critically about assessment with emerging bilingual students and how to take action on these issues. I am invested in finding avenues through which teachers can act to effectively improve assessment practices with emerging bilingual students.

Finally, I am a middle class, white female with intermediate Spanish language skills and with a college major in the French language, which makes me quite an outsider to the vast
majority of emerging bilingual students and parents in the United States. Thus, I had to make a concerted effort to continue to reflect upon and discuss aspects of this study with colleagues, teachers, and students whose background more closely mirrors the experience of emerging bilingual students in order to ensure that my conceptions of equity in assessment with linguistically and culturally diverse learners are not erroneous or misguided.

My background is similar, though, to the profile of many teachers. In the process of interviewing teachers, who potentially viewed me as an insider to the profession, teachers sometimes assumed that I understood the context of their interview responses. Thus, I needed to encourage teachers at times to be more specific in their interview responses so that they would provide contextual information that is important to the current study.

**Instrument Development and Data Sources**

This study employed two main data collection instruments with which to gather information about teachers’ critical use of assessment: a survey and two original interview protocols. Additionally, the study incorporated data from observation field notes and artifacts collected during the course of the research.

I generated interview prompts and survey items based on my conceptual framework (Figure 2). I developed the interview questions for the first interview protocol by creating items representing eight activities corresponding to each of the three critical thinking components, with 24 items in all. Table 4 shows each item and its corresponding activity and critical thinking component (System I, System II, and CE component). The only activities from the conceptual framework that were not included in the interview protocol were ‘copying,’ ‘comparing,’ and ‘contrasting.’ This was due to the fact that, occasionally, items representing a certain activity needed to appear twice in order to cover both language- and content-oriented concepts. Most of
the activities were represented in the items; however, and there were a consistent number of items across all three critical thinking components in the interview protocol.

Table 4

*Interview Protocol 1 Items and Their Corresponding Critical Thinking Activities and Components*

<table>
<thead>
<tr>
<th>Item</th>
<th>Component</th>
<th>Activity</th>
<th>Item Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>II</td>
<td>evaluate / judge</td>
<td>In any given unit, how do you know if all of your students are learning and making progress towards content area goals?</td>
</tr>
<tr>
<td>2</td>
<td>II</td>
<td>evaluate / judge</td>
<td>In any given unit, how do you know if all of your students are learning the language of the relevant content area?</td>
</tr>
<tr>
<td>3</td>
<td>II</td>
<td>evaluate / judge</td>
<td>How do you determine students’ grades at the end of a marking period?</td>
</tr>
<tr>
<td>4</td>
<td>I</td>
<td>know / understand</td>
<td>What do you think of when you think of the word “assessment”?</td>
</tr>
<tr>
<td>5</td>
<td>I</td>
<td>know / understand</td>
<td>What does the term “English Language Learner” or “ELL” mean to you? / What do you think of when you hear the term &quot;ELL”?</td>
</tr>
<tr>
<td>6</td>
<td>I</td>
<td>use / apply</td>
<td>What goals or objectives do you have for your students when you assess their progress?</td>
</tr>
<tr>
<td>7</td>
<td>II</td>
<td>integrate</td>
<td>Do you integrate lang development and content area learning goals in your assessments that you have in your classroom?</td>
</tr>
<tr>
<td>8</td>
<td>CE</td>
<td>create</td>
<td>When you devote time to formal assessment in your classroom, approximately what percentage of this time are the assessments self-created?</td>
</tr>
<tr>
<td>9</td>
<td>CE</td>
<td>adapt / transform</td>
<td>When you devote time to externally developed formal assessments, approximately what percentage of this time do you change or modify them?</td>
</tr>
<tr>
<td>10</td>
<td>CE</td>
<td>organize</td>
<td>When you devote time to assessment in your classroom, what percentage of this time are your students active participants in the assessment process?</td>
</tr>
<tr>
<td>11</td>
<td>I</td>
<td>use / apply</td>
<td>When you devote time to assessment in your classroom, what percentage of this time do you give everyone the same classroom assessments, including ELLs?</td>
</tr>
<tr>
<td>12</td>
<td>I</td>
<td>become familiar / listen</td>
<td>To what extent do you familiarize yourself with new assessments that your students will be evaluated by?</td>
</tr>
</tbody>
</table>
I developed the survey items based on the lessons learned when asking teachers the interview questions. For example, I learned that asking teachers if they involve parents in the assessment process did not elicit useful information, since most focus teachers shared that they did not involve parents in substantial ways. However, asking teachers if they created their own assessments elicited useful information, so I chose to include it as an item on the survey. Before administering the survey to participating teachers in the current study, the survey underwent the following development process: the continued use of a construct map (i.e. the conceptual

<table>
<thead>
<tr>
<th>Item</th>
<th>Component</th>
<th>Activity</th>
<th>Item Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 II</td>
<td>reflect</td>
<td>Do you feel prepared to design high-quality assessments that help you to understand how all of your students are progressing with the language and content in a unit of study?</td>
<td></td>
</tr>
<tr>
<td>14 II</td>
<td>reflect</td>
<td>Do you regularly set aside time to reflect on your assessment practices?</td>
<td></td>
</tr>
<tr>
<td>15 I</td>
<td>assume</td>
<td>Do you feel there is a characteristic way of engaging in assessment practices in this school?</td>
<td></td>
</tr>
<tr>
<td>16 I</td>
<td>comply</td>
<td>Does the administration ask that you conduct some assessment practices in a certain way?</td>
<td></td>
</tr>
<tr>
<td>17 I</td>
<td>become familiar / listen</td>
<td>Do you know what assessment methods other teachers in your school use?</td>
<td></td>
</tr>
<tr>
<td>18 CE</td>
<td>collaborate</td>
<td>When you plan your assessment methods, what percentage of this time do you and your colleagues collaborate?</td>
<td></td>
</tr>
<tr>
<td>19 CE</td>
<td>collaborate</td>
<td>Do you communicate with other teachers or specialists on a regular basis?</td>
<td></td>
</tr>
<tr>
<td>20 CE</td>
<td>collaborate</td>
<td>Do you involve parents in the assessment process in any way?</td>
<td></td>
</tr>
<tr>
<td>21 CE</td>
<td>adapt / transform</td>
<td>Do large-scale tests, such as TCAP, influence your use of classroom assessment in any way?</td>
<td></td>
</tr>
<tr>
<td>22 II</td>
<td>conclude</td>
<td>At what point do you think emerging bilingual students are no longer impacted by their language proficiency when taking assessments?</td>
<td></td>
</tr>
<tr>
<td>23 II</td>
<td>analyze / interpret</td>
<td>If you had the power to change anything about the assessments used with ELLs (either large-scale assessments or assessments in your school), would you change anything? If so, what would you change?</td>
<td></td>
</tr>
</tbody>
</table>
framework shown in Figure 2) to guide the revision of the questions, alignment of each item in the survey to a part of the conceptual framework, cognitive interviews with former and current teachers of emerging bilingual students, and a review by experts in the fields of assessment and culturally and linguistically diverse pedagogy.

This development procedure helped to ensure proper measurement of the construct throughout the process of creating the survey, administering it, and interpreting the survey responses. It is important to state clearly that it was impossible to create a survey or instrument protocol that can control for the natural human reaction to overstate one’s skills. Since it is socially desirable to appear highly capable in all endeavors, it is likely that all participants provided responses that painted a more positive portrait of their skill set than might a more objective observer. This is a limitation of both the survey and the interview protocols. As a single observer, I also might not have provided a completely objective picture of the six focus teachers’ skills during my observations of their lessons due to the fact that I could not hire a team of collaborators for this project with whom I could develop a reliable system of reporting across observers.

**Survey**

The survey included 17 items intended only to gather demographic information from respondents. There were 19 close-ended items and 4 open-ended items in the survey. The open-ended items were largely followed by the close-ended items, so as to allow teachers to be more specific with their answers and to ensure they are not led to an answer by various close-ended options. There were 13 Likert scale statements with six positive points, ranging from one to six. This range was selected so that respondents would not be able to easily identify the median
value, and so it would drive respondents to select a value that represents at least a mild opinion, instead of selecting a value which represents a neutral opinion.

Appendix A shows the survey in its final Qualtrics format. The survey was created using Qualtrics software. The survey was also delivered in Qualtrics format, which allowed for respondents to take the survey more easily. The components of thinking and action that are targeted in each item are notated on the survey.

*Interview Protocols*

Two interview protocols were used in this study. The first interview protocol was employed to gain insight into teachers’ critical thinking about and critical use of classroom and large-scale assessment with emerging bilingual students. The first interview protocol is presented in Appendix B. I used a semi-structured interview protocol using informal and clinical interviewing methods (McCracken, 1988; Rueda & Garcia, 1996). Research suggests that semi-structured interview questions are an ideal way to gather relatively personal knowledge, such as critical reflection about assessment, because stories and incidents are the communicative media that lend themselves best to express these thoughts and experiences. From these stories and incidents, one can infer knowledge and beliefs (Rueda & Garcia, 1996; Smith & Shepard, 1988). I began with questions that were open to teachers’ interpretations and questions that asked for their definitions of central concepts so that I could elicit a more valid picture of participants’ actual critical use of assessment.

I did not provide more information about the study to the participants than was necessary, so that I could minimize the extent of skewed responses when participants may choose to omit their real beliefs and instead share what they think is a more socially acceptable response. The first interview protocol covered topics regarding: (1) teacher background information, (2)
teachers’ broad conceptions of assessment, (3) defining terms, (4) teachers’ assessment practices, (5) the school and the community, and (6) the larger assessment environment. The teachers’ background information was recorded using Qualtrics survey software instead of being recorded and transcribed.

The second interview protocol was used to enable teachers to share their assessment practices as they use them in their own classrooms. The second interview protocol is presented in Appendix C. The interview focused on a recently administered and teacher-reviewed classroom assessment. Each teacher was asked to explain the entire procedure involved in the assessment process for that particular classroom assessment, from its inception to the consequences that follow the administration of the assessment. The final versions of the interview protocols are consistent with all the components of the conceptual framework (Figure 2).

Observation Field Notes

The three classroom observations that I conducted with each of the six focus teachers produced data in the form of field notes. I asked teachers, prior to observing their classrooms, which days they anticipated they would be engaged in assessment-related activities of some kind in order to ensure that I observed these events in the classroom. The classroom observations took place, then, during classroom periods which had been identified as “assessment-related.” My field notes were written in real time, which allowed me to more accurately capture the events that transpired during the course of the class period than taking the notes after the lesson had ended. The intent was to access teachers’ actual assessment practices with emerging bilingual students. I then used this information to help determine if teachers’ depictions of their assessment methodology during the interview sessions matched the practices they carry out in
the classroom. Thus, these observations served as a validation process of the interviews with teachers.

Artifacts

The artifacts that I collected primarily focused on assessment activities. In the second interview session, the teachers chose an assessment they implemented recently and they discussed the various aspects of the assessment in the context of the de-identified artifacts. I stored digital copies of the assessment artifacts for future reference. I also collected a copy of any worksheets or other artifacts used over the course of the classroom observations.

Recruitment and Data Collection Procedure

For the teachers who were working toward a CLD endorsement through a Master’s program or who already possessed such an endorsement by attending a Master’s program, I consulted the EECD program at the University of Colorado at Boulder for participants in my study. I asked the professors who taught an EECD Master’s class in the fall semester of 2014 to allow me to take a few minutes of their class time in order to introduce the survey to their classes. This was done with the intent to incentivize students to take the survey more due to the fact that they met the person who hosted the survey. There was also the additional incentive of several $100 prizes, which were awarded to three randomly selected participants. These awards were distributed in the spring of 2015, through the use of a random number generator in Excel to select the winners. The students were able to take the survey at any time, since the abbreviated web link (accomplished through the use of the web service “tiny url”) to the survey was distributed to students in the form of a paper handout. The survey was programmed to regretfully inform participants who were not currently teaching emerging bilingual students at least half-time, who had less than one year of experience in the teaching profession, or who were teaching
outside of the state of Colorado, that their participation was not being sought after at this time. This strategy ensured that the survey only collected data from participants who were part of the population of interest for the current study.

Since fewer than 30 current EECD Master’s students who fit the selection criteria responded to the survey from the classes, I elicited responses from former EECD Master’s students and teachers who received their CLD endorsement through educational programs other than the University of Colorado at Boulder. However, I began my search for survey participants within the population of current EECD Master’s students because they recently made the decision to learn more about pedagogy for emerging bilingual students. This decision to return to school was likely to signal one of two possibilities. Either there was a need for the teacher to learn about pedagogy for emerging bilingual students because this population was growing in their school, or the teacher took action to improve their practice for emerging bilingual students, and this action indicates that they have a greater tendency to use their agency as practitioners. Additionally, these teachers were exposed to issues of language in teaching in that current semester. Thus, issues of language and assessment with emerging bilingual students were more likely to be current in the minds of these teachers.

In order to recruit survey participants from the population of teachers who do not have and are not seeking a CLD endorsement through an educational program, I consulted the list of teachers who graduated from the undergraduate teaching program at the University of Colorado once my IRB protocol was approved. The staff in the School of Education at the University of Colorado sent out my survey request through email to initial licensure candidates that finished the program in 2010-11, 2011-12 or 2012-13. The email was sent to alumni from all licensure areas: Elementary Education, Mathematics, Science, English Language Arts, Social Studies,
Foreign Language, and K-12 Music. Again, I elicited responses only from teachers who had been teaching for at least one year, to ensure that teachers had adequate experience in their school environment to be able to respond with consideration to the questions in the survey. The recruitment methods across different years of graduation were also intended to ensure that the population of teachers who are Master’s students would not differ significantly by age and experience from the population of teachers who are alumnae of the undergraduate teacher education program, since Master’s students tend to be older and more experienced in their profession.

Since I did not obtain at least 30 respondents for the survey from each CLD status group from these initial recruiting methods, I needed to search for additional survey participants using new recruiting methods. In order to encounter more survey participants, I tried various strategies for recruitment. First, I attended two local educational conferences to ask teachers in attendance to participate in my survey. I provided a physical handout for teachers to take with them that included both the link to the survey and a quick-response, or QR, code that allowed users to simply take a picture of the square image on the paper using their smartphone in order to access the link. The conferences I attended in order to recruit survey participants included the Colorado Association for Bilingual Education (CABE) Conference called *Equity: Growing and Leading Together*, hosted by CABE on September 25, 2014. This conference yielded survey participants who had or were working towards a CLD endorsement through an educational program. I also attended the 2014 Colorado Science Conference on November 21, 2014 at the Denver Merchandise Mart, which was hosted by the Colorado Association of Science Teachers (CAST). This conference yielded mostly survey participants who did not possess a CLD endorsement.
In order to further recruit survey participants, I also asked people who work in the field of education to forward my survey to current teachers of emerging bilinguals in Colorado. These people included fellow doctoral students and one former doctoral student of the EECD program at the University of Colorado at Boulder, a fellow doctoral student in the education department at the University of Colorado at Denver, a friend who works at a charter school network in Denver, teachers with whom I worked on previous research projects in education, as well as fellow teachers with whom I took Master’s level classes at the University of Colorado at Boulder. Although the strategy of convenience sampling for survey participants was not ideal, it was a necessary approach to take in order to attain the required number of participants (n=30) in each category of CLD status.

When I contacted potential survey participants for recruiting purposes through email, I included a short description of the survey and its estimated length along with the link to the survey on Qualtrics, which was the server that housed the survey online. Teachers signed the consent form online as the first page of the survey, and the survey did not advance to the next page unless the survey participant agreed to the terms of the consent form. The survey also ended promptly if the participant indicated that they were not currently teaching in a public, charter, or private school in Colorado, or if they did not have at least one emerging bilingual student in their classes.

The surveys were administered after interviewing six initial participants, the “focus” teachers, and observing their classrooms. This was different than the original strategy proposed in the prospectus. The order of instrument implementation was changed due to the fact that I realized the survey would be better informed by the interview process, since there would be more context provided from the interviews to better develop the survey. If the survey had been
conducted first, there would have been little opportunity to understand how teachers interpreted
the survey items before using those items to inform the interview questions for teachers, since
many of the survey items were close-ended questions with little room for commentary.
Additionally, it became more apparent that the concept of selecting a sub-group of participants
(n=6) for the interviewing and observation process from the larger pool of surveyed participants
(n=60) was theoretically desirable, yet practically unrealistic. A range of skill levels of critical
assessment use for the interview and observation process was nevertheless accomplished through
sampling teachers from various backgrounds, as is further discussed in the results and analysis
section of this study report.

The data collection procedure also involved interviews and observations with a set of six
“focus” teachers and only interviews with an additional four teachers in the study. Two interview
sessions were conducted with each of the ten teachers, both in a semi-structured interview
format. The first interview protocol focused on the teacher’s ideas regarding assessment
practices with emerging bilingual students and their ideas about these practices in the
environment around them. The second interview protocol focused on the teacher’s assessment
example and how she/he put assessment ideas into practice. These interview protocols were
always conducted during multiple sessions, mostly due to scheduling constraints but also to
prevent interview fatigue. In some cases, it took as many as five separate sessions to complete all
of the interview questions in the two protocols. Teachers were given time between the
administration of the first and second interview protocols to gather assessment artifacts for the
second interview protocol. The interviews also served to clarify any events in the observations
that I found confusing or that I thought I might have misinterpreted.
Additionally, I conducted three observations in each of six focus teachers’ classrooms in order to gain insight into the context of each teacher’s practice. I refrained from taking an active or participatory role in the classroom. The observations were conducted when the teacher indicated they would be engaged in assessment-related activities, though these activities were in the context of any content area the teacher chose. It was ideal when, at times, the teacher administered an assessment and subsequently we discussed that assessment in the second interview protocol. That was not always possible due to scheduling constraints; however, the teacher always described the assessment artifact selected for the second interview in detail so I could understand it fully even if I was not present at the time of its administration.

**Data Analysis**

In order to address all of the research questions, I transcribed all the first interviews with the six focus teachers that I interviewed and observed. I transcribed the first interviews with the additional four teachers I interviewed; however, I had sufficient data from the interviews and observations with the focus teachers such that this data did not need to be part of the analysis. The second interview with all interviewed teachers was not part of the analysis for the same reason.

Then, I organized each of the six focus teachers’ responses by item, so that I could see all teachers’ responses to a single item together, instead of reading through and analyzing one teacher’s interview without considering how other teachers responded to the questions. This helped me to ensure that I was as impartial as possible when analyzing and coding the teachers’ interview responses. Then, I coded the data using a combination of process and provisional coding techniques.
Coding of the Interviews

Process coding involves the sole use of gerunds to signify action in the data set, according to Saldaña (2009) in his portrayal of the method based on the work of Charmaz (2002). Both observable activities (such as using, copying, etc.) and conceptual activities (such as analyzing, reflecting, etc.) can be coded using a process coding strategy (Saldaña, 2009). Process in data is “represented by sequences of action/interaction/emotions changing in response to sets of circumstances, events, or situations” (Corbin & Strauss, 2008, p. 98). The content of the data, combined with one’s interpretation of this content, determines how process is conceptualized.

Provisional coding, according to Saldaña (2009), in summarizing the work of Miles and Huberman (1994), establishes a predetermined “start list’ set of codes prior to fieldwork” (Miles and Huberman, 1994, p. 58). The provisional list of codes can be generated from the study’s conceptual framework and research questions, which is precisely the technique applied in the current study. The current coding system was generated from the study’s conceptual framework, using each activity listed in the conceptual framework as a different code used to analyze the data. I used provisional coding combined with process coding by using only gerunds, yet limiting myself to specific predetermined sets of gerunds as represented in the conceptual framework of the current study. Thus, I coded activities to better reveal processes that occurred in the interviewed discussion, but these activities were chosen from a set of nine activities each from System I, System II, or the CE component in the conceptual framework. This strategy ensured that the codes would speak to the research questions and test the conceptual framework as it was designed for the current study.
I coded all of the interview responses according to one master procedure. This master procedure first necessitated that I develop codes to fit all instances of teachers’ responses. In order to firmly align the codes to the research questions, the codes chosen directly mirror the conceptual framework used in Figure 2. I read through all the responses and decided where independent statements were made in each response. Independent statements were chosen based on the continuity of purpose in a statement. If a statement could stand alone as a single idea, it was chosen as an independent statement. If there was more than one idea being expressed within a certain response, it was divided into different statements. Even if some statements were related to each other, I chose to keep statements separate if they could stand alone in purpose.

Then, for every statement in the responses from each interview question from each teacher, I decided if the statement represented an activity from System I, System II, the CE component, or if it did not represent any component from the conceptual framework in Figure 2 at all. In order to make this coding process transparent, I developed a code book to help me and to help others see the criteria used to decide to which component a statement belonged. The code book is shown in Appendix D. For each component, the code book shows a table containing the nine activities in the component, the general definition of each activity, a contextualized definition of each activity, a hypothetical example of each activity with respect specifically to assessment with emerging bilingual students, and an example from the current study’s interview data to demonstrate how the codes were actually applied to the data. Some of the activities are listed in tandem with another activity. For example, in System I, “listen” is listed together with “become familiar.” Although the activities listed together did not always occur simultaneously in the interview data, they were related to one another and therefore they could co-occur. For example, “listening” and “becoming familiar” are both activities that a teacher can engage in to
learn about what her students know. At other times, a teacher might become familiar with what her students know through reading or noticing physical changes in the student’s behavior. This involves a teacher becoming familiar with what her students know without also being involved in the activity of listening. For this reason, paired activities were coded together when it was relevant to do so; otherwise they were coded separately.

Independent statements were occasionally divided in the process of coding. The only instances where independent statements were divided involved situations where an independent statement depicted more than one activity from System I, System II, or the CE component. If, for instance, a teacher discussed both ‘listening’ to her students and promptly ‘adapting’ her lesson based on that information, then the statement was spliced at the point where the first activity ended and the second activity was taken up in the discussion. This strategy was undertaken in order to prevent the assignment of multiple labels of activities from System I, System II, and the CE component to a single independent statement.

Once the interview data were coded with appropriate labels on each independent statement for all activities including System I, System II, and the CE component, the frequency of each coded activity was tallied for each teacher’s response to each interview question. These frequencies were added together to obtain the total frequency of each coded activity in each component for each teacher. Then, I was able to calculate, for each teacher, the percentage she/he discussed a particular component-related activity relative to the total instances of component-related activities that emerged in the interview.

Then, I developed a narrative of each teacher to describe her/his use of critical thinking and action in her/his practice relevant to assessment for emerging bilinguals as depicted by what she/he discussed in her/his interview. These narratives represent a case study approach to
describing the interview data, allowing for a rich, contextualized understanding of the teacher’s approach to critical thinking and action. In order to accomplish this, I analyzed the use of the various component-related activities in two different, but complementary, ways. I analyzed the teachers’ use of critical thinking and action using the quantitative information from the calculations of the code frequencies from each teacher to help me attain a global perspective of the activities they used more frequently and those that they rarely, if ever, used in the interview sessions. I also made use of the qualitative information from the coded responses for each activity to develop a more specific sense of how the teachers used these component-related activities in the context of assessing emerging bilingual students. The narratives combine each of these data sources to paint a picture of the teachers’ use of critical thinking and action with both broad and narrow brush strokes.

Coding of the Observations

The observations were coded in a method that was almost identical to that of the interview coding method. However, there were a few important variations in the way the two types of data were coded, due to the fact that the nature of the observation data is different from that of the interview data. First, before engaging in the coding method used in the interviews, the observation data underwent a preliminary coding process. This was done to ensure that all the observation data would be relevant to the current study. The observation field notes included many interactions between the teacher and her/his students, so it was helpful to first identify the parts of each observed class period that were relevant to the areas of either second language acquisition or of assessment. In order to identify the relevant parts of the class periods more clearly, a table was created to organize the information from each observed class period. The rows in the table kept a record of each distinct activity that took place during the observed class
period. Columns in the table kept a record of the following: the start time of the activity, a
description of the class activity, the role of the teacher, the role of the students, instruction in the
activity that is helpful for emerging bilingual students, and informal or formal assessment
practices that took place during the activity. An example of this table can be found in Table 5.

Table 5

Sample of the Preliminary Coding Method for the Classroom Observation Data

<table>
<thead>
<tr>
<th>Time of Activity (Start Time)</th>
<th>Description of Activity</th>
<th>Role of the Teacher</th>
<th>Role of the Students</th>
<th>Instruction Helpful for Emerging Bilinguals</th>
<th>Informal or Formal Assessment Practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:05</td>
<td>They begin reading</td>
<td>Storyteller</td>
<td>Listeners of the story</td>
<td>Listening to oral language, comparing vocabulary words in native language: “A serial is like a story that keeps going, like a <em>novela</em>. It’s a story that keeps going and going and going.”</td>
<td>She pauses in the story and asks if students know what certain words are, or if they understand certain ideas.</td>
</tr>
</tbody>
</table>

*Note.* This table was created based on the field notes from Valera’s first classroom observation.

The preliminary coding method of creating tables for each classroom observation was completed within two days after the observations were conducted in order to ensure that the classroom activities were accurately organized and summarized. After I developed a coding system and a code book to guide the coding process for the interviews, I applied the same coding system to the organized tables of observation data. Some rules were followed for the observation coding process that were distinct from the interview coding process.
First, the only component-related activities that were paired for the observation coding were ‘adapting’ and ‘transforming.’ All other component-related activities were coded as single activities, since the activities were more concretely identifiable within the context of the observations. For example, it was clear when a teacher ‘listened’ to students during an observation. When a teacher ‘became familiar’ with students’ understanding, it was not through the oral modality of language. Instead, the teacher ‘became familiar’ with students’ understanding when she used a kinesthetic or written method of collecting information about learners’ content knowledge. When teachers discussed this topic in the interviews, it was sometimes difficult to tell if this process occurred through oral, written, or kinesthetic interactions. Thus, the two activities would be paired together so as to prevent an imprecise assignment of codes to the data.

Second, the System I activity of ‘using’ was applied to the observation data only when certain resources were employed. The classroom whiteboard, notebooks, and paper were all considered to be so common to a classroom that they were not counted as a resource that a teacher might ‘use’ as a component-related activity. Computer programs, movies, digital media, novels, and manipulatives were all considered to be resources that a teacher might ‘use’ as a component-related activity to assist with their instruction.

Third, due to the fact that the interviews had more instances of described activity, the interviews had many more instances of System I, System II, and CE component activity codes. Despite the smaller sample size of codes, the coded observations were still illustrative because they provided a different source of data to increase the validity of the findings for the current study.
Addressing the Research Questions

In order to answer the first research question (*To what extent do teachers of emerging bilingual students think critically about assessment?*), I examined the System II codes in the observation and interview data of the six focus teachers, as these are the codes that are relevant to the first research question. I examined the survey responses for System II activities as well. I used the coding results to group the teachers from least to most critical. I was able to identify specific activities that set some teachers apart from others in their ability to think critically with regard to assessment for emerging bilingual students. These activities were used to create a profile of the characteristics of teachers who think about assessment critically with their emerging bilingual students.

To address the second research question (*To what extent do teachers of emerging bilingual students use assessment critically?*), I examined the CE component codes in the observation and interview data of the six focus teachers, as these are the codes that are relevant to the second research question. I used the coding results to group the teachers in order from least to most critical. I was able to identify specific activities that set some teachers apart from others in their ability to act critically with regard to assessment for emerging bilingual students. I examined the survey responses for CE component activities as well. There is a considerable overlap between the way in which I addressed teachers’ critical thinking for Research Question 1, and the way in which I addressed teachers’ activities that are critically oriented in Research Question 2. This is to be expected, as thinking and action are inter-connected (Burbules & Berk, 1999). However, I used the categories of CE component-related activities to help me focus more on purposeful action as opposed to thought. The observation field notes and associated artifacts
were particularly important sources of data to help address the second research question because they revealed teachers’ real assessment practices in their classrooms.

To address the third research question (Which factors influence teachers’ critical thinking about and critical use of assessment for emerging bilingual students?), I examined the codes across teachers from the activities pertaining to System I, as this is the component that was most directly related to this research question. I also analyzed the survey results that corresponded to System I activities. I used these results to identify external and internal factors that can shape critical thinking and action among teachers. The former are outside factors that positively or negatively impact teachers, such as an administration that requires their teachers to use specific assessments. The latter are attitudes and beliefs that influence teachers’ willingness and motivation to act to change a given situation. The intent was to identify the internal and external factors more highly associated with lower or higher levels of critical thinking. Also, I analyzed broader characteristics to identify if a teacher’s formal educational experience might influence their critical use of assessment. The responses to the survey were categorized into two groups of 30 teachers each according to formal CLD educational experience. This categorization facilitated the analysis of the CLD educational experience as a variable affecting a teacher’s critical use of assessment.
Chapter 5
Findings

Teacher Narratives

For each focus teacher that participated in both the interviews and observations, a narrative was written to describe the demographics of the teacher as well as the kinds of component-related activities in which they participated. All names used are pseudonyms, which were used to protect the privacy of the teachers. The narratives depict the teachers’ relative levels of engagement in critical activities. Specific frequency descriptors were used to uniformly characterize the relative proportion of component-related activities in each of the narratives. The frequency descriptors correspond to a particular range of percentages. This correspondence is depicted in Table 6. The following narratives are included for each teacher, with the interview portion first and the observation portion second. The narratives provide an overview of the component-related activities employed by each teacher. The narratives are elaborated to a much greater extent in Appendix E.

Table 6

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Frequency descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>00.00-10%</td>
<td>Hardly any of / none of (for 0%)</td>
</tr>
<tr>
<td>10.01-20%</td>
<td>A few of / a small percentage of</td>
</tr>
<tr>
<td>20.01-40%</td>
<td>Some of</td>
</tr>
<tr>
<td>40.01-60%</td>
<td>Approximately half of</td>
</tr>
<tr>
<td>60.01-80%</td>
<td>Most of</td>
</tr>
<tr>
<td>80.01-90%</td>
<td>Nearly all of</td>
</tr>
<tr>
<td>90.01-100%</td>
<td>Practically all / all of (for 100%)</td>
</tr>
</tbody>
</table>
Adele

*Interview Narrative.* The first teacher who participated in the interview is Adele. Adele teaches 5th grade at a public school. Approximately 26-50% of students in her class were classified as “English Language Learners” through the school identification process. In the 2013-2014 academic year, 24.3% of students in this school were identified as ELLs and 59.8% of students were eligible for free or reduced lunch (Public School District, 2015). She did not have a CLD endorsement through an educational program at the time of the interviews. She was planning to attend two classes that summer of 2014 through the Public School District’s Professional Development organization to attain the CLD endorsement. It was the last month of her first year of teaching at the time of the interviews, and her first year of teaching as an elementary generalist. This teacher was in the 26-35 year age range, and completed her teacher education at a local university. She has a certification in elementary education. The school in which Adele taught did not yet have a bilingual program, but was identified as a Transitional Native Language Instruction (TNLI) school and was transitioning to a bilingual model the following academic year. She speaks English natively and speaks American Sign Language as a second language at a beginning level.

Adele incorporated evidence of System II thinking most into her discussion of self-identified behavior during the interview, with some (37.62%) component-related discussion belonging to System II. She demonstrated the least evidence of CE component-related activities in her discussion during the interview, with some (28.71%) component-related discussion belonging to the CE component. Some (33.66%) component-related discussion was identified as System I related. Overall, an analysis of interactions with this teacher revealed that she has a
capacity for CE component activities but does not engage in these activities the majority of the time. These proportions are visually represented in Figure 4.

![Figure 4](image)

*Figure 4.* The proportion of component-related activities Adele employed in interview sessions.

*Observation Narrative.* With respect to the observations, Adele incorporated evidence of System I thinking most into her activities during class, with most (62.96%) of the component-related activity belonging to System I. She demonstrated the least evidence of System II activities during the observations, with hardly any (7.41%) of the component-related activity belonging to System II. Some (29.63%) of the component-related activity was identified as CE component-related. These proportions are visually represented in Figure 5.
Alma

*Interview Narrative.* The second teacher who participated in the interview is Alma. Alma teaches 4th grade at a public school, which is the same school in which Adele teaches. Approximately 51 - 75% of her students over the past three years are classified as “English Language Learners” through the school identification process. In the 2013-2014 academic year, 24.3% of students in this school were identified as ELLs and 59.8% of students were eligible for free or reduced lunch (Public School District, 2015). She was working towards a CLD endorsement through an educational program at the time of the interviews. She was attending a local university to attain this CLD educational experience. It was her 10th year of teaching, and she taught this grade for 6-9 years. This teacher is in the 36-45 year age range, and completed her teacher education in elementary education. She has a teaching certification in Colorado as an elementary generalist. She also holds the CLD Bilingual endorsement. The school in which Alma teaches did not yet have a bilingual program, but was identified as a Transitional Native...
Language Instruction (TNLI) school and was transitioning to a bilingual model the following academic year. She is Latina, speaks Spanish natively, and speaks English as a second language fluently and practically natively.

Alma incorporated evidence of CE component activities most into her discussion of self-identified behavior during the interview, with approximately half (48.15%) of component-related discussion belonging to the CE component. She demonstrated the least evidence of System I activities in her discussion during the interview, with some (24.44%) of the component-related discussion belonging to System I. Some (27.41%) of the component-related discussion was identified as related to System II. Overall, this teacher revealed that her critical thinking and action process mostly involved CE component activities due to the relative proportion of time spent adapting and transforming her practice based on her students’ needs. These proportions are visually represented in Figure 6.

Figure 6. The proportion of component-related activities Alma employed in interview sessions.
Observation Narrative. With respect to the observations, Alma incorporated evidence of System I processes most into her activities during class, with approximately half (58.33%) of component-related activity belonging to System I. She demonstrated the least evidence of System II activities during the observations, with a small percentage (16.67%) of component-related activity belonging to System II. Some (25.00%) of the component-related activity during class was identified as CE component-related. These proportions are visually represented in Figure 7.

![Figure 7](image)

*Figure 7.* The proportion of component-related activity Alma employed in classroom observation periods.

Valera

Interview Narrative. The third teacher who participated in the interview is Valera. Valera teaches English Language Arts and English Language Development (ELD) at the fifth grade level at a public school. She ‘platoons’ with her teaching partner, which means that she teaches English Language Arts, while her partner teaches Mathematics, Social Studies, and Science. Approximately 75-100% of her students over the past three years are classified as “English
Language Learners” through the school identification process. At the school level in the 2013-14 academic year, 75.6% of students were classified as ELLs and 98.8% were eligible for free or reduced lunch (Public School District, 2015). She has a CLD endorsement through an educational program. She attended a local university to attain this CLD experience. It is her ninth year of teaching, and she completed 4-5 years of teaching these subjects. This teacher is in the 36-45 year age range. She has a teaching certification in Colorado as an elementary generalist and was just finishing working towards a bilingual CLD endorsement at the time of the interview. The school in which Valera teaches has an early exit transitional bilingual program. The students in her classroom, then, only receive an ELD block during their instructional week due to the fact that they are in upper elementary and have already transitioned out of Spanish language instruction. Valera is Latina and speaks English natively, though some Spanish was spoken at home when she was a child. She speaks Spanish as a second language at a high intermediate level.

Valera incorporated evidence of CE component thinking most into her discussion of self-identified behavior during the interview, with approximately half (45.36%) of the component-related discussion belonging to the CE component. She demonstrated the least evidence of System I activities in her discussion during the interview, with some (22.68%) of the component-related discussion belonging to System I. Some (31.96%) of the component-related discussion was identified as System II related. Overall, this teacher revealed that she engages in CE component-related activities most due to the way she adapts her assessment process for emerging bilinguals. These proportions are visually represented in Figure 8.
Figure 8. The proportion of component-related activity Valera employed in interview sessions.

Observation Narrative. With respect to the observations, Valera incorporated evidence of CE component thinking most into her activities during class, with approximately half (53.57%) of component-related activity belonging to the CE component. She demonstrated the least evidence of System II activities during the observations, with hardly any (7.14%) component-related activity belonging to System I. Some (39.29%) of the component-related activity during class was identified as System I related. These proportions are visually represented in Figure 9.
Figure 9. The proportion of component-related activity Valera employed in classroom observation periods.

**Lucas**

*Interview Narrative.* The fourth teacher who participated in the interview is Lucas. Lucas teaches fifth grade Mathematics, Science, and Social Studies in the same public school as Valera does. Approximately 75-100% of his students over the past three years are classified as “English Language Learners” through the school identification process. At the school level in the 2013-14 academic year, 75.6% of students were classified as ELLs and 98.8% were eligible for free or reduced lunch (Public School District, 2015). He does not have a CLD endorsement through an educational program. It is his third year of teaching, and his first year of teaching these specific subjects at the fifth grade level. This teacher is in the 26-35 year age range, and completed his teacher education in a regular teacher education program in the Midwest region of the United States. He has a certification to teach as an Elementary Generalist and as an Early Childhood Generalist. The school in which Lucas teaches has an early exit transitional bilingual program. He speaks English natively and does not speak another language.
Lucas incorporated evidence of System I thinking most into his discussion of self-identified behavior during the interview, with approximately half (40.86%) of the component-related discussion belonging to System I. He demonstrated the least evidence of System II activities in his discussion during the interview, with some (25.81%) of the component-related discussion belonging to System II. Some (33.33%) of the component-related discussion was identified as CE component-related. These proportions are visually represented in Figure 10.

![Figure 10](image)

**Figure 10.** The proportion of component-related activity Lucas employed in interview sessions.

*Observation Narrative.* With respect to the observations, Lucas incorporated evidence of System I thinking most into his activities during class, with most (63.16%) of the component-related activity belonging to System I. He demonstrated the least evidence of System II activities during the observations, with hardly any (5.26%) component-related activity belonging to System II. Some (31.58%) of the component-related activity during class was identified as CE component-related. These proportions are visually represented in Figure 11.
Figure 11. The proportion of component-related activity Lucas employed in classroom observation periods.

Serena

*Interview Narrative.* The fifth teacher who participated in the interview is Serena. Serena teaches sixth, seventh, and eighth grade Mathematics at a charter school. Almost all (75-100%) of her students over the past three years are classified as “English Language Learners” through the school identification process. At the school level in the 2013-14 academic year, 77.6% of students were classified as ELLs and 95.6% were eligible for free or reduced lunch (Public School District, 2015). She had just completed an educational program to attain a CLD endorsement at the time of the interviews and observations. She attended a local university to attain this CLD experience. It is her fourth year of teaching, and her third year of teaching Mathematics. This teacher is in the 26-35 year age range, and completed her teacher education through the Teach for America program. She has a certification in Secondary Mathematics and in CLD education. The school in which Serena teaches does not have a bilingual program. They
do have most signs in the school in both Spanish and in English, though. She speaks English natively and speaks Spanish as a second language at an intermediate level.

Serena incorporated evidence of CE component thinking most into her discussion of self-identified behavior during the interview, with approximately half (43.92%) of component-related discussion belonging to the CE component. She demonstrated the least evidence of System I activities in her discussion during the interview, with some (21.62%) component-related discussion belonging to System I. Some (34.46%) component-related discussion was identified as System II related. Overall, this teacher revealed that she had much more CE component-related behavior, due to her self-identified emphasis on social justice in her teaching during the interviews. These proportions are visually represented in Figure 12.

![Pie Chart](image)

**Figure 12.** The proportion of component-related activity Serena employed in interview sessions.

**Observation Narrative.** With respect to the observations, Serena incorporated evidence of System I thinking most into her activities during class, with approximately half (47.37%) of component-related activity belonging to System I. She demonstrated the least evidence of
System II activities during the observations, with a small percentage (13.16%) of component-related activity belonging to System II. Some (39.47%) of the component-related activity during class was identified as CE component-related. These proportions are visually represented in Figure 13.

![Figure 13. The proportion of component-related activity Serena employed in classroom observation periods.](image)

**Claire**

*Interview Narrative.* The sixth teacher who participated in the interview is Claire. Claire teaches sixth grade Science at a charter school. Almost all (76-100%) of her students over the past three years are classified as “English Language Learners” through the school identification process. At the school level in the 2013-14 academic year, 78.7% of students were classified as ELLs and 94.6% were eligible for free or reduced lunch (Public School District, 2015). She holds a CLD endorsement through the state of Colorado, yet it was not received through the experience of an educational program. She passed the PLACE test in order to gain this teaching endorsement. It is her fifth year of teaching, and her fifth year of teaching Science. Claire is in
the 26-35 year age range, and she completed her teacher education through the Teach for America program. She has a state certification in Secondary Science and CLD education. The school in which Claire teaches does not have a bilingual program. She speaks English natively and speaks Spanish as a second language at a beginning level.

Claire incorporated evidence of System I thinking most into her discussion of self-identified behavior during the interview, with some (36.05%) component-related discussion belonging to System I. She demonstrated the least evidence of System II activities in her discussion during the interview, with some (30.23%) component-related discussion belonging to System II. Some (33.72%) component-related discussion was identified as CE component-related. Overall, Claire revealed that System I-type activity took slight prominence in her practice, though her activities related to all three components were similar in proportion. These proportions are visually represented in Figure 14.

![Figure 14](image.png)

*Figure 14.* The proportion of component-related activity Claire employed in interview sessions.

*Observation Narrative.* With respect to the observations, Claire incorporated evidence of System I thinking most into her activities during class, with approximately half (60.00%) of
component-related activity belonging to System I. She demonstrated the least evidence of System II activities during the observations, with a small percentage (16.00%) of component-related activity belonging to System II. Some (24.00%) component-related activity during class was identified as CE component-related. These proportions are visually represented in Figure 15.

![Figure 15](image)

*Figure 15.* The proportion of component-related activity Claire employed in classroom observation periods.

**Survey Results**

**System I**

Three items on the survey target System I concepts in teachers’ thoughts and actions. The first item asks teachers from which sources they receive the most helpful information about their ELLs’ progress towards content area goals. This item specifically identifies how teachers use and apply information from their classroom to help them in their content area assessment practice for emerging bilingual students. Table 7 shows teachers’ responses, disaggregated by CLD educational experience status. The majority (60%) of all teachers responded that in-class activities provided the most information about how their ELLs are progressing towards content
area goals. Seventeen percent of teachers thought that students’ questions or conversation in class provided the most helpful information, whereas 13% of teachers suggested that classroom tests yielded the most beneficial information. None of the teachers thought that they receive the most information from large-scale tests. This indicates that teachers thought the most helpful information to them was that which was collected within their own classroom, specifically informal assessment episodes, such as in-class activities. If a teacher’s administration does not value these same kinds of assessment types, it could serve as a barrier for the teacher to engage in System II processes. For example, if a teacher is asked to look at large-scale tests to analyze and interpret their ELLs’ progress, they will likely be unable to engage in this process as meaningfully as if they were also able to look at other data sources to achieve this goal.

Table 7

Response to “I receive the most helpful information about my ELLs’ progress towards CONTENT AREA goals from:” Percentages

<table>
<thead>
<tr>
<th></th>
<th>Culturally and Linguistically Diverse (CLD) Experience</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teachers who have CLD educational program experience</td>
<td>Teachers who do not have CLD educational program experience</td>
<td>Both CLD and non-CLD experienced teachers</td>
</tr>
<tr>
<td>Homework</td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Classroom tests</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Quizzes</td>
<td>10</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>In-class activities</td>
<td>63</td>
<td>57</td>
<td>60</td>
</tr>
<tr>
<td>Students’ questions / conversation in class</td>
<td>13</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>Large-scale tests</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

The second item asks teachers from which sources they receive the most helpful information about their ELLs’ progress towards language goals. This item specifically identifies how teachers use and apply information from their classroom to help them in their assessment of
language for emerging bilingual students. Table 8 shows teachers’ responses, disaggregated by CLD educational program status. The majority (58%) of all teachers responded that in-class activities provided the most information about how their ELLs are progressing towards language goals. Thirty-three percent of respondents thought that students’ questions or conversation in class provided the most helpful information instead. Only 2% of teachers thought that they receive the most information from large-scale tests. This indicates that teachers thought the most helpful information to them was that which was collected within their own classroom, whether in regards to language or content area goals. This shows that teachers readily find informal assessment information from their classrooms helpful and therefore they may use this data in ways that are potentially critical, if they also incorporate System II and the CE component when analyzing assessment information. Yet, outside sources of data and even formal classroom assessment data by themselves may be less helpful for teachers, which makes it more difficult for teachers to use this data in a potentially critical manner.

Table 8

*Response to “I receive the most helpful information about my ELLs' progress towards LANGUAGE PROFICIENCY goals from: ” Percentages*

<table>
<thead>
<tr>
<th></th>
<th>Culturally and Linguistically Diverse (CLD) Experience</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teachers who have CLD educational program experience</td>
<td>Teachers who do not have CLD educational program experience</td>
<td>Both CLD and non-CLD experienced teachers</td>
</tr>
<tr>
<td>Homework</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Classroom tests</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Quizzes</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>In-class activities</td>
<td>60</td>
<td>57</td>
<td>58</td>
</tr>
<tr>
<td>Students' questions / conversation in class</td>
<td>30</td>
<td>37</td>
<td>33</td>
</tr>
<tr>
<td>Large-scale tests</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
The third item asks teachers how prepared and confident they feel in designing and implementing high-quality assessment methods to measure how their ELLs are progressing with a unit of study. This item specifically identifies the level of knowledge teachers think they possess with language assessment methods for emerging bilingual students. Table 9 shows teachers’ responses, disaggregated by CLD educational program status. Almost half (45%) of all teachers responded that they feel mostly prepared and confident with regards to assessment methods for ELLs. 33% of respondents felt somewhat prepared and confident, whereas only 10% felt very prepared. A problematic finding is that 12% of teachers thought that they could benefit from the help of a class or professional development on this topic. Although proportionally more teachers without CLD program experience felt unprepared to design and implement assessment for ELLs, teachers from both groups shared this sentiment. This indicates that some teachers struggle with System I aspects of their teaching practice, which limits their ability to move to System II- and CE component-type processes. Thus, teachers could readily benefit from further education on the topic.

Table 9

Response to “When it comes to designing and implementing high-quality assessment methods to measure how my ELLs are progressing with a unit of study, I feel:” Percentages

<table>
<thead>
<tr>
<th>Culturally and Linguistically Diverse (CLD) Experience</th>
<th>Teachers who have CLD educational program experience</th>
<th>Teachers who do not have CLD educational program experience</th>
<th>Both CLD and non-CLD experienced teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very prepared and confident</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Mostly prepared and confident</td>
<td>57</td>
<td>33</td>
<td>45</td>
</tr>
<tr>
<td>Somewhat prepared and confident</td>
<td>27</td>
<td>40</td>
<td>33</td>
</tr>
<tr>
<td>I could really use the help of a class or professional development for this</td>
<td>7</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
System II

Four items on the survey target System II concepts in teachers’ thoughts and actions. The first item asks teachers how they use helpful information about their ELLs’ progress in a class. This item specifically aims to understand how teachers integrate information from their classroom to help them decide on next steps with their emerging bilingual students. Table 10 shows teachers’ responses, disaggregated by CLD educational program status. Interestingly, there were no clear majorities of responses among all teachers. Rather, 37% of all teachers indicated that they use information to make inferences to inform their ELLs’ content area instruction, 25% indicated that they use information to make inferences to inform their ELLs’ language instruction, and 25% suggested that they use information to get a good idea of where students are in their learning without using it to inform future instruction.

A greater proportion of CLD experienced teachers (33% vs. 17%) explained that they make inferences to inform future language instruction, whereas a greater proportion of non-CLD experienced teachers (47% vs. 27%) indicated that they make inferences to inform future content area instruction. Non-CLD experienced teachers (33%) also were more likely to suggest that they use information to understand where students are in their learning without using it to inform their future instruction in any way. This indicates that teachers mostly use information to inform their future instruction either with regard to content (37%) or language (25%). This is an encouraging result, because using information to inform future instruction is emblematic of a more critical use of assessment than simply taking note of the information or using it mostly to give grades. Next, the focus needs to centered on developing teachers’ awareness of aspects of language in adapting their future instruction.
Table 10

Response to “When I receive helpful information about my ELLs’ progress in a class, I MOSTLY:” Percentages

<table>
<thead>
<tr>
<th></th>
<th>Culturally and Linguistically Diverse (CLD) Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teachers who have CLD educational program experience</td>
</tr>
<tr>
<td>Use it to help me in my grading and reporting of student progress</td>
<td>7</td>
</tr>
<tr>
<td>Make inferences from it to help me know what to do for my students in future language instruction</td>
<td>33</td>
</tr>
<tr>
<td>Make inferences from it to help me know what to do for my students in future content area instruction</td>
<td>27</td>
</tr>
<tr>
<td>Use it in some other way</td>
<td>13</td>
</tr>
<tr>
<td>Use it to get a good idea of where students are in their learning</td>
<td>17</td>
</tr>
<tr>
<td>Non-response</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
</tr>
</tbody>
</table>

The second item asks teachers how they conceptualize assessment in their minds. This item specifically aims to understand the conclusions teachers have made about the practice of assessment. Table 11 shows teachers’ responses, disaggregated by CLD educational program status. A slim majority of all teachers (52%) concluded that assessment is part of their daily practice. Additionally, 20% of all teachers responded that assessment allows them to learn more about their students but they struggle to know how to use that information. Thirteen percent of all teachers indicated that assessment is a practice that is part of their unit-level instruction. This small segment of teachers still use assessment to drive their instruction in broad scope, but not nearly as consistently as the 52% of teachers who incorporate it into their daily practice.
It should be noted that proportionally more CLD-experienced teachers selected that assessment is part of their daily practice (60% vs. 43%), whereas proportionally more non-CLD-experienced teachers selected that they benefit from assessment even though it’s hard to know how to use the information gathered from it (27% vs. 13%). This suggests that CLD-experienced teachers may be more adept than their non-CLD-experienced counterparts at using assessment information to make decisions to inform their instruction. Using assessment to drive instruction represents a more critical use of assessment, thus these results indicate that CLD-experienced teachers may be more equipped to use assessment critically. These results also show that almost half of the respondents could stand to learn how to use assessment more critically in order to drive their language- and content-oriented instruction on a daily basis.

Table 11

Response to “I MOSTLY think assessment is:” Percentages

<table>
<thead>
<tr>
<th>Culturally and Linguistically Diverse (CLD) Experience</th>
<th>Teachers who have CLD educational program experience</th>
<th>Teachers who do not have CLD educational program experience</th>
<th>Both CLD and non-CLD experienced teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>A practice that is part of my daily instruction</td>
<td>60</td>
<td>43</td>
<td>52</td>
</tr>
<tr>
<td>A practice that takes time away from my instruction</td>
<td>3</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>A way to learn more about my students and what they can do, but it's hard to know how to use that information</td>
<td>13</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td>Something else</td>
<td>13</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>A practice that is part of my instruction on a unit level</td>
<td>10</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

The third and fourth item in System II both center on teachers’ interpretation of how important classroom and large-scale assessments are to various stakeholders: teachers
themselves, their students, their students’ parents, their administration, and their district. The responses are organized to show their responses in regards to both classroom and large-scale assessments, according to the type of stakeholder. Figure 16 shows the importance of these assessments to teachers themselves. The mean value of importance ascribed to classroom assessments by teachers is 5.30, whereas the mean value of importance given to large-scale assessments by teachers is 3.00 (where 6.00 is always important and 1.00 is never important). This shows that, on average, teachers value classroom assessment more than large scale assessment. In fact, no teachers responded that classroom assessment was rarely or never important to them, and no teachers responded that large-scale assessments were always important to them. This indicates that teachers find classroom assessment more helpful to their System II processes than large-scale assessments.

![Figure 16](image)

*Figure 16. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to me, as a teacher” percentages.*
Figure 17 shows the importance of these assessments to the students, as interpreted by teachers. The mean value of importance ascribed to classroom assessments is 4.47, whereas the mean value of importance given to large-scale assessments is 2.90 (where 6.00 is always important and 1.00 is never important). This shows that, on average, teachers think their students value classroom assessment more than large-scale assessment as well. In fact, no teachers responded that classroom assessment was never important to their students, whereas only one teacher responded that large-scale assessments were always important to their students. This indicates that teachers interpret their students as more likely to find classroom assessment more helpful to them than large-scale assessments.

![Figure 17](image)

*Figure 17. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to my students” percentages.*

Figure 18 shows the importance of these assessments to the students’ parents, as interpreted by teachers. The mean value of importance ascribed to classroom assessments is 4.17,
whereas the mean value of importance given to large-scale assessments is 3.37 (where 6.00 is always important and 1.00 is never important). This shows that, on average, teachers think their students’ parents value classroom assessment more than large-scale assessment as well. This indicates that teachers interpret their students’ parents as more likely to find classroom assessment results more helpful to them than large-scale assessment results.

![Figure 18. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to my students’ parents” percentages.](image)

Figure 18 shows the importance of these assessments to respondents’ principals and administration, in their interpretation. The mean value of importance ascribed to classroom assessments is 4.47, whereas the mean value of importance given to large-scale assessments is 4.83 (where 6.00 is always important and 1.00 is never important). This shows that, on average, teachers think their principals and administration value classroom assessment similarly to large-scale assessments, yet they ascribe slightly more importance to large-scale assessment. This
indicates that teachers interpret their principals and administration as a little more likely to find large-scale assessment results more helpful to them than classroom assessment results.

![Figure 19. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to my principal and administration” percentages.](image)

Figure 20 shows the importance of these assessments to the respondents’ district. The mean value of importance ascribed to classroom assessments is 4.27, whereas the mean value of importance given to large-scale assessment is 5.50 (where 6.00 is always important and 1.00 is never important). This shows that, on average, teachers think their district values classroom assessment less than large-scale assessment. This indicates that teachers interpret their district as more likely to find large-scale assessment results more helpful to them than classroom assessment results.
Figure 20. Response to “In my interpretation, I think assessments (classroom and large-scale) are important to my district” percentages.

CE Component

Five items on the survey target CE component concepts in teachers’ thoughts and actions. The first item asks teachers if they create their own assessments. This item specifically identifies how teachers create information-gathering tools to help them in their assessment practice for emerging bilingual students. Table 12 shows teachers’ responses, disaggregated by CLD educational program status. There was no clear trend that represented the responses of most teachers. The majority (73%) of teachers responded that they create their own assessments usually, frequently, or always. Only 3% of teachers responded that they rarely or never create their own assessments. Proportionally more non-CLD-experienced teachers create their own assessments always or frequently, whereas proportionally more CLD-experienced teachers usually, sometimes, or rarely create their own assessments. This indicates that most teachers at least usually create their own assessments, which is representative of teachers engaging their CE component processes. However, the results show that proportionally fewer CLD-experienced
teachers create their own assessments with as much frequency, which indicates that they are less engaged in CE component processes with regard to this aspect.

Table 12

*Response to “In general, I create my own assessments” Percentages.*

<table>
<thead>
<tr>
<th>Culturally and Linguistically Diverse (CLD) Experience</th>
<th>Teachers who have CLD educational program experience</th>
<th>Teachers who do not have CLD educational program experience</th>
<th>Both CLD and non-CLD experienced teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always – 6</td>
<td>10</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>Frequently - 5</td>
<td>20</td>
<td>37</td>
<td>28</td>
</tr>
<tr>
<td>Usually – 4</td>
<td>33</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td>Sometimes - 3</td>
<td>33</td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>Rarely – 2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Never – 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

The second item asks teachers if they change or modify parts of assessments that they do not create themselves. This item specifically identifies how teachers change or modify information-gathering tools to help them in their assessment practice for emerging bilingual students. Table 13 shows teachers’ responses, disaggregated by CLD educational program status. As with the item regarding teachers’ creation of assessment tools, there was no clear trend that represented the response of most teachers. The majority (69%) of teachers responded that they change or modify parts of assessments sometimes, usually, or frequently. Proportionally more non-CLD-experienced teachers change or modify assessments always or frequently. This indicates that most teachers change or modify assessments sometimes, usually, or frequently, which is representative of teachers engaging their CE component processes. However, the results show that proportionally fewer CLD-experienced teachers change or modify assessments with as
much frequency, which indicates that they are less engaged in CE component processes in this respect.

Table 13

Response to “I change or modify parts of assessments that I do not create myself” Percentages

<table>
<thead>
<tr>
<th>Culturally and Linguistically Diverse (CLD) Experience</th>
<th>Teachers who have CLD educational program experience</th>
<th>Teachers who do not have CLD educational program experience</th>
<th>Both CLD and non-CLD experienced teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always - 6</td>
<td>10</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>Frequently - 5</td>
<td>20</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>Usually - 4</td>
<td>23</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>Sometimes - 3</td>
<td>23</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>Rarely – 2</td>
<td>10</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Never – 1</td>
<td>13</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

The third item asks teachers if they collaborate with their colleagues to plan their assessment methods. This item specifically identifies how teachers actively engage CE component processes by working together to facilitate their assessment practice. Table 14 shows teachers’ responses, disaggregated by CLD educational program status. Approximately one third of respondents frequently collaborate with colleagues, and another third of respondents sometimes collaborate with colleagues. Only 9% of teachers responded that they rarely or never collaborate with their colleagues. Proportionally more non-CLD-experienced teachers collaborate with colleagues always or frequently, whereas proportionally more CLD-experienced teachers collaborate sometimes or rarely. These results indicate that most teachers collaborate with their colleagues at least sometimes, which represents teachers engaging their CE component processes. However, the results show that proportionally fewer CLD-experienced teachers
collaborate with as much frequency, which indicates that they are less engaged in CE component processes in this respect.

Table 14

Response to “My colleagues and I collaborate to plan our assessment methods” Percentages

<table>
<thead>
<tr>
<th>Culturally and Linguistically Diverse (CLD) Experience</th>
<th>Teachers who have CLD educational program experience</th>
<th>Teachers who do not have CLD educational program experience</th>
<th>Both CLD and non-CLD experienced teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always – 6</td>
<td>10</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>Frequently - 5</td>
<td>27</td>
<td>43</td>
<td>35</td>
</tr>
<tr>
<td>Usually – 4</td>
<td>7</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Sometimes - 3</td>
<td>43</td>
<td>23</td>
<td>33</td>
</tr>
<tr>
<td>Rarely – 2</td>
<td>13</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Never – 1</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>No response</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

The fourth item asks teachers if they would change anything about assessment used with ELLs if they had the power. All teachers responded that they would change something about assessment used with ELLs if they could. The item then asked what the participants would change, given the choice. This item specifically identifies how teachers actively engage in CE component processes by proposing and imagining alternatives and solutions to current problems with assessment of emerging bilingual students. Figure 21 shows all teachers’ coded responses, disaggregated by CLD educational program status, to the question: “What would you change, and why?” The coded responses are shown in percentages, with the total representing the sum of the surveyed teachers’ coded responses, not the sum of individual teachers’ responses. This is important to distinguish because sometimes an individual teacher responded such that their answer was categorized into two or more separate codes. The most frequently occurring coded response (other than miscellaneous responses or teachers who responded that they weren’t sure)
was to employ native language assessment (n=14), with a slightly higher proportion of non-
CLD-experienced teachers who provided this response. The next most frequently occurring
coded response was to make language more accessible when testing content (n=12). Slightly
higher proportions of CLD-experienced teachers gave this response. The third most frequently
occurring coded response was a suggestion for an elimination or reduction of testing, or simply
to exclude ELLs from such testing (n=11). This coded response was more commonly observed in
CLD-experienced teachers. The fourth most frequently cited coded response was a proposal to
make assessments more culturally relevant or responsive (n=8). Slightly more CLD-experienced
teachers gave this response than non-CLD-experienced teachers.
Figure 21. Coded responses from both CLD-endorsed and non-CLD-endorsed teachers to “If you could change anything about assessment used with ELLs (either large-scale assessments or classroom-based assessments), would you?” All respondents said yes and then were prompted to respond to: “What would you change, and why?” percentages.
The fifth item asks teachers what holds them back from making the changes they suggested in the previous item. This item specifically targets why teachers may be able to engage in CE component processes enough to propose or imagine alternatives and solutions to issues surrounding assessment with emerging bilinguals, but still may not be able to organize or create change themselves. This item essentially identifies potential barriers to teachers’ abilities to take further action. Figure 22 depicts teachers’ coded responses to this question, disaggregated by CLD-educational program status. The coded responses are shown in percentages, with the total representing the sum of the surveyed teachers’ coded responses, not the sum of individual teachers’ responses. The most common coded response (n=19) to this question pertained to requirements at the school, district, state, or federal level acting as a barrier for teachers to make changes. The second most frequent coded response (n=14) pertained to teachers’ sense of power and control over the situation. The next most common coded response (n=12) given by respondents was the demand of time on their ability to orchestrate change. Similar proportions of CLD-experienced teachers as non-CLD experienced teachers gave responses of requirements as well as power and control. More CLD-experienced teachers cited time as a barrier to their ability to make changes, while more non-CLD-experienced teachers named resources as a barrier. This provides evidence to suggest that teachers of all backgrounds may be activating their CE component processes through imagining alternative solutions to problems they witness, yet they are held back from making changes by similar forces. To help teachers reach their CE component potential, some factors may need to be placed more in teachers’ control.
Figure 22. Responses of teachers with and without CLD educational experience to the question: “What holds you back from making changes?” percentages.

Research Question One: To What Extent do Teachers of Emerging Bilingual Students Think Critically About Assessment?

Focus Teachers’ System II Activities and Related Factors Influencing Their Critical Use of Assessment

The focus teacher who had the highest proportion of System II activities in the results from the interview and observation data is most engaged in critical thinking with respect to assessment for emerging bilingual students. Based on the interview data, Adele was reported to have the highest percentage of System II activities, followed by Serena and Valera. Based on the observation data, Alma was reported to have the highest percentage of System II activities, closely followed by Claire and Serena. With respect to the combined interview and observation...
data, Serena was reported to have the greatest percentage of System II activities, closely followed by Claire and Adele. No single teacher clearly outpaced another in this category. Rather, the differences among them were minor, since the teacher who had the highest percentage of System II thinking varied depending on the data source. Table 15 shows the percentage of component-related activity recorded for all six focus teachers according to the various data sources.

Table 15

<table>
<thead>
<tr>
<th></th>
<th>Interview</th>
<th>Observation</th>
<th>Mean from both data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adele</td>
<td>33.66</td>
<td>28.71</td>
<td>48.31</td>
</tr>
<tr>
<td></td>
<td>37.62</td>
<td>62.96</td>
<td>22.52</td>
</tr>
<tr>
<td>Alma</td>
<td>24.44</td>
<td>48.15</td>
<td>41.39</td>
</tr>
<tr>
<td></td>
<td>27.41</td>
<td>58.33</td>
<td>22.04</td>
</tr>
<tr>
<td>Valera</td>
<td>22.68</td>
<td>45.36</td>
<td>30.98</td>
</tr>
<tr>
<td></td>
<td>31.96</td>
<td>39.29</td>
<td>19.55</td>
</tr>
<tr>
<td>Serena</td>
<td>21.62</td>
<td>43.92</td>
<td>34.50</td>
</tr>
<tr>
<td></td>
<td>34.46</td>
<td>47.37</td>
<td>23.81</td>
</tr>
<tr>
<td>Lucas</td>
<td>40.86</td>
<td>33.33</td>
<td>52.01</td>
</tr>
<tr>
<td></td>
<td>25.81</td>
<td>63.16</td>
<td>15.53</td>
</tr>
<tr>
<td>Claire</td>
<td>36.05</td>
<td>33.72</td>
<td>48.02</td>
</tr>
<tr>
<td></td>
<td>30.23</td>
<td>60.00</td>
<td>23.12</td>
</tr>
</tbody>
</table>

Differences in System II Across Data Sources

After examining Table 15, it became clear that there are substantial differences in the percentage of System II activities in the interviews relative to the observations. In fact, each teacher was recorded to have at least 10 fewer percentage points in System II activities from the observations as compared to System II activities from the interviews. The relatively low incidence of System II activity in the observations is complemented by a relatively high incidence of System I activities in the observations, across all teachers. Conversely, there is a
relatively low incidence of System I activities and a relatively high incidence of System II activities in the interviews. Figure 23, Figure 24, and Figure 25 show the focus teachers’ relative engagement in the three component-related activities based on the interview data, the observation data, and the mean of both data sources, respectively. These figures facilitate the observation of the differences across the interview and observation data, as well as differences among the focus teachers’ activities.

![Figure 23](image)

*Figure 23.* Focus teachers’ relative engagement in System I, System II, and CE component activities from the interview data in percentages.
This difference could be attributed to a couple of factors. First, it is human nature to slightly overstate one’s skills in an interview situation because it is socially desirable to respond in this fashion. This phenomenon is called “overclaiming” (Atir, Rosenzweig, & Dunning, 2015). Since I asked teachers to self-assess their ability to thoroughly consider issues of language
in their assessment practice for their emerging bilingual students, it is likely that teachers felt social pressure to respond in a way that makes their practice appear highly skilled. This phenomenon of overclaiming is human nature, and it is also likely that the participants overstated their skills in somewhat equal proportions. People are more likely to overclaim “to the extent that they perceive their personal expertise favorably” (Atir et al., 2015, p. 1295). Therefore, the teachers who thought themselves to be more skilled in the area of assessment for emerging bilingual students were more slightly more likely to have actually overstated their skills.

Of the focus teachers, Alma and Lucas both thought they did not feel prepared to design high-quality assessments for their emerging bilingual learners. Valera, Serena, and Claire all felt as though they had become better prepared to design assessments for their language learners over time. Rather ironically, the only first year teacher in the study was also the only teacher who communicated that she felt prepared in this endeavor. She acknowledged that it was a bold statement for a first year teacher, but with the focus on data-driven instruction she noted that she had learned a lot already and felt prepared in that area. Therefore, based on the available research, Adele would be most likely to overstate her skills, whereas Alma and Lucas would be least likely to overstate their skills. However, it is impossible to definitely tell if and to what extent participants overestimated their skills, so no claims have been made based upon this theory.

Second, it is to be expected that one would observe a relatively higher proportion of System II related activities in the context of an interview and a relatively higher proportion of System I related activities in an observation. This is due to the fact that System II related activities, such as reflecting or analyzing, take place mostly in one’s own mind and are often unobservable. System II activities emerge naturally when an interviewer asks questions that
beckon the interviewee to employ these activities. Teachers sometimes discuss ideal practice in interviews since they share how they would teach in an optimal situation, though it is sometimes difficult to carry this out every day in reality. System I related activities are often observable, since they include actions such as listening or using resources. These activities emerge more naturally in the context of a lesson observation, since the observer can readily witness the teacher engaging in these activities.

System II activities were the least common type of component-related activity in the observation data. System II activities were also the least frequently observed type of component-related activity for all teachers except for Adele. It is possible that Adele may have been particularly engaged in System II related activities because it was the end of her first year of teaching, so she reflected rather extensively on this new practice during the interview. It is to be expected that System II activities are less frequently observed in all people, not only the six focus teachers in this study. The finding that the focus teachers employed less System II activities than other component-related activities is consistent with Kahneman’s theory on how people use System I and II in everyday situations. In his book, he shows that System II is lethargic, and people typically employ System I thinking most in everyday life. The relative magnitude of the differences between teachers’ System I and System II use is less important than their relative differences. This suggests that there are areas in which teachers can improve with regard to thinking critically about assessment for emerging bilingual students.

**Survey Findings: Some Teachers Could Use Support on Using Assessment to Inform Their Instruction**

There were a couple of interesting items on the survey that elicited information about how teachers use assessment for emerging bilingual students. When asked how they use helpful information about their ELLs’ progress in a class, 25% of teachers responded that they use it to
get a good idea of where students are in their learning and 5% of teachers responded that they use it to help them with grading and reporting their students’ progress. Essentially, 30% of teachers do not use it to inform their future instruction. However, 62% of teachers did select that they use helpful information to make inferences about their future language or content area instruction. Thus, the majority of survey participants are critical users of assessment information when they receive it, though approximately a third of participants could learn more about how to use assessment information to impact their subsequent instruction.

This claim is also supported by the findings from the survey item that asked teachers what they thought assessment was mostly about. One fifth of the survey respondents selected that assessment is a “way to learn more about my students and what they can do, but it's hard to know how to use that information.” While most teachers (52%) viewed assessment as a part of their daily instructional practice, it is significant that one in five survey participants admitted that they find it challenging to know how to inform their instruction with assessment information they gather from students. Some teachers may benefit from professional development or other sources of support to help them learn to integrate assessment information into their reflection and planning for subsequent lessons.

**Areas to Improve: Metacognitive Awareness of Language in Assessment**

Specifically, one area in which teachers might further improve their critical skills is to develop metacognitive awareness of the role language plays in assessment. Valera demonstrated how one might engage in this practice when she posed metacognitive questions to her students about the language of the assessment. She incorporated this practice when at least some students did not demonstrate mastery on a test item. She did this in order to determine which aspects of the language were challenging. Valera used this information to analyze how to optimally phrase
or structure the language in her assessments. However, she is the only teacher that engaged in
this practice. The other teachers in the study did not ask metacognitive questions about the
language of their assessments, though two focus teachers (Serena and Claire) also posed
metacognitive questions to their students about assessment in general.

All teachers could benefit from using Valera’s approach to assessment analysis: she
collects information about how students interpreted the language on a challenging assessment
item and then determines if any linguistic aspects of her assessment should be modified the next
time she assesses that standard. Since she expects students to develop English academic
language, she retains items that contain challenging language terms or structures on her
assessment while adding an additional item to assess the targeted content of an item using
modified language. In this manner, she can isolate the targeted content more directly in one item
and assess the targeted content using its associated academic language in a different item. This
allows her to distinguish if students are struggling with the linguistic- or the content-based aspect
of the item. Teachers may better target the source of students’ challenges if they can identify the
source of the issues with more certainty. This type of analysis is a System II related activity that
teachers might employ to help serve their emerging bilinguals in assessment situations. In order
to engage in this type of practice, teachers might also need the training to recognize how to
modify the language of an assessment. They may also benefit from additional time allocated to
assessment planning in order to carry out any necessary modifications.

*Areas to Improve: Representativeness Heuristic in Decision-Making*

Another area in which teachers might improve their critical thinking and action skills is
with regard to avoiding the heuristic of representativeness. In his book, “Thinking fast and
slow,” Kahneman (2010) discusses an idea of “representativeness” that people use in their
decision-making processes (p. 149). Representativeness is a heuristic that people employ when they focus on the similarity of a description of a person to a stereotyped case, while ignoring the proportion of a certain case relative to all cases in the population. Kahneman’s idea of “representativeness” reveals that teachers who are not using their System II-type activities to guide their thought processes may jump to conclusions about the kind of student that an emerging bilingual will be in their class, based on information about their language status. The realities are that each student learns in a unique manner, based on a plethora of both external and internal influences. Although learning a second language in school adds a tremendous amount of new content that students must learn, second language learners tend to make greater academic gains as a population in Colorado, relative to comparable students in their schools and school districts (Escamilla et al., 2005). The idea that there is an achievement gap caused by emerging bilingual student status in Colorado is not supported by evidence. Escamilla et al. (2005) also identify a deficit view expressed by some professionals that emerging bilinguals are expected to fall behind in school due to language issues. They describe the “language as a problem” orientation, where people view language as a negative obstacle to students’ success instead of as an asset.

Teachers may fall into a trap when they use the representativeness heuristic to develop expectations of how their students will perform in their class. As is clear in the literature on the subject, when teachers have high expectations of their students, their students make more gains than when teachers have lower expectations of them (Rosenthal & Jacobson, 1963). Teachers’ expectations become a self-fulfilling prophecy. This makes teachers’ expectations, and the heuristics that help them to arrive at their expectations, an important factor contributing to students’ academic success.
Some teachers fell prey to the heuristic trap of “representativeness” when they were surveyed for the study. When asked what they would change about assessment for emerging bilingual students, three percent of teachers’ coded responses communicated a deficit perspective of emerging bilingual students. For example, one teacher claimed that, “They [ELLs] will all do poorly on PARCC because of limited language ability” (Non-CLD surveyed teacher 9, survey data, October 16, 2014). This teacher’s expectations of his students on content-area assessments are low due to the fact that he connects his students’ performance to his students’ language status. As described above, this can lead to a self-fulfilling prophecy over the course of the school year. It is important to note that none of the CLD-educated teachers communicated a deficit perspective. All three of the teachers who shared this view did not have educational experience through a CLD program.

Some teachers also fell prey to the heuristic trap of “representativeness” when they were interviewed for the study. For example, when asked what they thought of when they thought of the word “English Language Learner,” Claire mentioned that “we have 98% ELL designation or very close to that, so I typically think of my students, who tend to come in pretty low to 6th grade, but are usually on track for grade level by the end of their 7th grade year” (Claire, personal communication, May 5, 2014). Claire did not recognize that she fell a little into the representativeness heuristic trap in developing her expectations of her students around their language status. She directly associated students’ ELL designation to their level of performance in content area knowledge when they entered her classroom. This is a powerful stereotype, because she was able to overcome the representativeness heuristic with regard to other similar factors – such as language variety. She noted: “I definitely know I have a bias towards thinking of ELLs as being Spanish speakers because of the neighborhood we’re in” (Claire, personal
communication, May 5, 2014). This reveals that the stereotype of ELL status as a negative contributing factor in teachers’ expectations can be powerful, since Claire was able to overcome bias due to the representativeness heuristic with regard to language variety but not with regard to language status as a contributing factor to academic success or failure.

Valera showed an understanding of the implicit connection that can be made about language status and its impact on academic success or failure. She expressed that, when asked what she thinks of when she hears the term “ELL,” she thinks of “minority. Um. And that bothers me because I think that implies a deficit” (Valera, personal communication, March 28, 2014). She described how stereotypes affect the way teachers in her school think about their emerging bilinguals and their expectations of them. Thus, she shows a higher level of System II-type activities in her thought processes regarding emerging bilingual students and expectations of their performance. Unfortunately, Valera came to the conclusion that this understanding must be “my own baggage” because of her background growing up as a Latina. When I suggested that it might be society’s baggage, she retorted that “I still see it with the teachers I work with” (Valera, personal communication, March 28, 2014). While most people develop initial impressions based upon stereotypes, either consciously or unconsciously, it is misguided to make professional judgments based on such stereotypes.

Valera illustrates how teachers might overcome making such misguided judgments by avoiding the representativeness heuristic and instead viewing language as an asset to students. Other teachers were less skilled in using System II type thinking when making these judgments. For example, Adele expressed “I think being able to speak another language is great, but I can’t imagine what it would have been like if he never had to, um, if he never had a second language” (Adele, personal communication, May 15, 2014). Thus, Adele could benefit from learning how
to incorporate more System II thinking when she makes professional judgments about her emerging bilinguals. Valera summarized these issues well when she noted:

[I]n my heart I agree with that definition [of ELLs as students qualified for language services], but the realities of society don’t allow me to fully mesh that definition with what it truly means. Like I know that, if you were to ask the majority of people in a safe and secure environment what they truly believe ‘ELL’ meant, that’s not what people would come up with. They would come up with deficit notions, like ‘we have to jump through hoops to teach these kids,’ ‘teach “those” kids,’ because those are the words that they use. (Valera, personal communication, March 28, 2014)

Conclusion

In examining the extent to which teachers of emerging bilingual students think critically about assessment, it became clear that none of the focus teachers stood out as engaging in critical thinking considerably more than any other. There were differences in the extent to which teachers engaged in System II, or critical thinking related, activities based on the type of data source. Teachers expressed more System II activities in interviews than they did in observations, likely due to the context in which the data was collected and as a result of the factor of social desirability to express System II activities. The survey data revealed that a majority of teachers think critically about assessment information to help inform their future instruction for emerging bilingual students. However, most teachers did not state that they used assessment information to guide their future language instruction for language learners. Instead, teachers focused on the use of assessment information to guide future content area instruction. Furthermore, a small yet significant percentage of teachers did not know how to use assessment information to guide their future instruction and admitted that they find it challenging to do so. This is an area in which
some teachers could benefit from support. It also became apparent that teachers could benefit from professional support in learning how to develop metacognitive awareness about the role language plays in assessment. Finally, teachers could benefit from support to help them avoid the representativeness heuristic in developing expectations about their emerging bilingual students.

Research Question Two: To What Extent do Teachers of Emerging Bilingual Students Use Assessment Critically?

Focus Teachers’ CE Component Activities and Related Factors Influencing Their Critical Use of Assessment

The focus teacher who had the highest percentage of CE component activities in the interview and observation data engaged the most in critical action with respect to assessment for emerging bilingual students. Based on the interview data, Alma was reported to have the highest percentage of CE component activities, followed by Valera and Serena. Based on the observation data, Valera was reported to have the highest percentage of CE component activities, followed by Serena and Lucas. With respect to the combined interview and observation data, Valera again emerged with the highest percentage of reported CE component activities, followed by Serena and Alma.

Valera

According to the results, then, Valera appears to have the highest percentage of CE component activities. Valera, Serena and Alma together comprise the group of teachers with above-average CE component skills. This finding accords with the qualitative results provided in the narratives and the holistic assessment of the researcher. Valera was not only reported to have the highest proportion of CE component-related activities, she often took action despite the fact that her critical activities were not encouraged by her school’s culture. For instance, Valera
would reach out to families of students whose standardized test scores did not match her assessment of the student’s progress in the class in order to alert them to the fact that their grading system was mandated by the school and to make suggestions for the family to increase their chances of gaining access to rigorous middle schools despite the official record of lower scores in elementary school. Since attending graduate school Valera admitted that she would “get into a lot more trouble,” by pointing out weaknesses in district-sponsored assessments to district personnel (Valera, personal communication, April 9, 2014). This put Valera’s relationship with the district personnel, and potentially even her job security, at risk in order to strive for a fairer assessment environment for her emerging bilingual students. Valera even left her position at the end of the academic year in which I collected data in order to pursue a coaching position for emerging bilingual students, where she could exert more influence on assessment and instructional practices than in her individual classroom. Teachers who engage in critical activity, despite the fact that it is not encouraged, show authentic desire to be critical users of assessment.

Additionally, although this did not surface from the quantitative analysis of the codes, Valera actively encouraged her students to become critical thinkers in their own lives. This type of activity should have been assigned an individual code, in retrospect. Her activities were labeled as CE component activities, yet the activity did not accurately capture the complexity of this type of interaction. Often Valera’s encouragement served a social justice purpose, as it is an act of social justice to ask students from traditionally subordinated backgrounds to be critical of the sociopolitical context that fosters this subordination. Valera would ask her students to critically examine their situation. For example, when one of Valera’s students exclaimed that she couldn’t learn anything from her classmates during a peer assessment writing activity, Valera engaged in a discussion about how her classmates all have valuable knowledge to share due to
their different life experiences (Valera, personal communication, April 9, 2014). Valera explained that she was not the ultimate source of knowledge simply because she was an authority figure as a teacher. This shows that Valera was able to use assessment in a way that helped to empower her students through peer assessment. Students could feel valued and capable when they were asked to provide an assessment of their own peer’s work. Valera also created a classroom environment in which her students engaged in a discussion and the teacher served as a mediator instead of the leader of the conversation. Students contributed their thoughts about whether or not they thought it was ever acceptable to lie based on the textual evidence from the book they read, *Esio Trot*, but also based on their life experience. Valera legitimized her students’ life experiences by encouraging them to share these experiences as part of a discussion to challenge them academically and socio-emotionally. In doing this, she enabled students to move from “object to subject position” to take meaningful part in the classroom community (Bartolome, 1994, p. 177).

Valera also posed metacognitive questions to her students, as witnessed in her observations and in her interview. Valera asked her students which items on her unit assessment they found to be challenging and why they might have chosen the incorrect answer. The class had a discussion about how some vocabulary words, in particular, were tricky for them. This kind of metacognitive discussion helps students to learn how to self-monitor their progress and to assess themselves. This fosters independence on the part of the students as learners. It also allows students to speak from their own vantage points, in order to “create learning contexts in which students are able to empower themselves throughout the strategic learning process” (Bartolome, 1994, p. 188). Thus, when teachers support students to engage in metacognitive processes to assess their own learning, it helps to create students who are critical thinkers. As
mentioned previously, teachers who inspire their students to become critical thinkers must be critical themselves first.

**Serena**

Serena was also reported to have a higher percentage of CE component activities than other teachers in the study. Like Valera, she actively supported her students to become critical thinkers in their own lives. Though this type of activity was not given the special attention it deserved in the coding process, Serena encouraged her students to engage in critical thinking for the purposes of social justice more than any other teacher in the study. She asked students to take part in project-based learning for authentic purposes, which were intended to instill feelings of success and confidence in students. When students saw that they were able to solve problems that serve a purpose in the world outside of school, they felt empowered and capable. As Lilia Bartolome explains, “creating democratic learning environments where students become accustomed to being treated as competent and able individuals” is important in order to prepare students to deal with larger society’s inequalities (Bartolome, 1994, p. 179). Serena also discussed how she wanted to move in a trajectory where her instruction and assessment would challenge students into questioning the unequal power structures that exist in society. For example, Serena described how she wanted to question students about:

> [W]hat percentage of the land around in our neighborhood has parks, what percentages, looking at the difference between percentages of liquor stores versus grocery stores in our neighborhoods, versus other neighborhoods in Denver, and what that looks like. Or, looking at income levels and trying to find ways to talk about bigger issues through the math. And then they actually get to practice, and ... So, that's where I want to be headed towards. (personal communication, May 16, 2014)
Serena felt that it was imperative to begin asking her sixth graders to question why our society is inequitable, particularly towards people from minority language backgrounds. Her goal was to bring more of this focus into the classroom so that students might learn to be critical of their situation. It is impossible to lead this kind of critical discussion in a classroom without first having a critical perspective on societal structure oneself. Therefore, Serena’s focus on helping her students learn to be critical was indicative of the type of critical thinker she was. It is clear that she also takes critical action based on her critical thinking, as she took action to help her own group of students also learn this skill.

Serena also engaged in these social justice-oriented initiatives in her classroom despite the fact that her students would never be assessed by the school-wide tests on these topics, and these initiatives were not recognized or encouraged by her administration. Serena therefore was authentic in the way she acted critically to engage her emerging bilingual students in thinking critically themselves. She carried out this project-based, social justice-oriented instruction without any particular incentive. In fact, she had some disincentive to promote critical thinking in her students because this concept was not part of the skill set assessed every six weeks that comprised students’ grades, which is how teachers are evaluated as well. Serena’s critical activities could then be characterized as authentic in nature and not caused by external forces.

Serena also encouraged metacognitive thinking processes in her students, to enable them to become more independent learners capable of assessing their own progress. She helped students to track their own learning to see how they had improved over the course of a semester. She asked students to keep a journal, called a “resource book,” of how students learned each content and language objective so that they could go back and refer to their journal later. She also asked students, “What helps you focus in math?” on an exit ticket during an observation.
Asking students to attend to these metacognitive processes helps students become more independent learners, and encourages them to be critical thinkers because they learn to be in control of their own learning and development.

*Alma*

Alma was also reported to have a high percentage of CE component activities in her interview and observations. Alma expressed goals similar to Serena’s in her desire for her teaching to be social justice-oriented and for students to learn to be critical thinkers as well. She admitted to having barriers from the imposed school curriculum to making her teaching social justice-oriented, since she talked about how she would “sneak in” reading about Civil Rights or “carve out time” for a read aloud about these topics (Alma, personal communication, April 18, 2014). However, Alma still incorporated ways to “talk in depth about those social justice kind of things so that education is still transformative in nature. So it’s not just, I’m gonna take this test” (Alma, personal communication, April 18, 2014). She discussed how she taught students the history behind the “n” word so that students would understand the context and hopefully decide on their own to avoid using the word. Instead of simply telling students not to use the word once she heard it in her class or punishing them for using the word, Alma helped give students background on the issue and asked students to use proper judgment. This approach develops critical thinking skills in her students, as she asks students to consider the wider socio-political context before making their own decisions. Alma also brought in “Teaching Tolerance” videos to directly teach students about some of the historically inequitable power structures that have caused struggle or conflict.

Alma related how the focus on testing can hamper the development of students’ critical thinking skills. She described how she knows that prompt writing is not the only kind of writing
that students will need to do in their lives. She was quite upset when she told her students that they were going to spend time writing and they asked her, “Well what’s the prompt?” (Alma, personal communication, April 18, 2014). She was adamant about wanting her students to think critically about life, and she went so far as to disagree with the instructional leaders of the district when they asserted that students simply need to focus on prompt writing. In this way, Alma also shows engagement in authentic critical activity because she is moved to act in a critical way despite the fact that it could threaten her job security or her relationship with district leaders.

Commonalities Among Teachers With Large Percentages of CE Component Activities

The characteristic of ‘authenticity’ is an important distinguishing factor that Valera, Serena, and Alma all share. Their critical activity can be juxtaposed against Claire’s critical activity, which was often catalyzed by mandated activity in her school. Claire transformed her instruction after extensive analysis of her assessments every six weeks, for example, but it was difficult to tell if she would engage in this transformation if it were not mandated by the administration.

The other distinguishing factor that Valera, Serena, and Alma share is that they all focused on developing critical thinking habits in their students. In order to take on this task properly, a teacher must first have a good understanding of what it means to be critical oneself. In this way, these three teachers exemplified engagement in critical activity by empowering their students who, as emerging bilinguals, often already arrive in their classrooms feeling disempowered by societal norms and expectations.

An interesting finding is that Valera, Serena, and Alma all shared CLD educational experience at the Master’s level. While Alma decided to discontinue the program at the approximate time of the data collection for this study, it is striking that all three of them had
completed at least one course at the University. From the group of teachers without CLD educational experience, Claire was certified in CLD education by the state of Colorado and Adele was about to begin her CLD certification process, but none of them had entered or planned to enter a Master’s program for CLD education.

**CE Component Survey Findings**

The survey results provided a broader perspective of how teachers use assessment for emerging bilingual students critically. However, the survey was best equipped to address narrow, specific ways in which teachers might use assessment critically. Whereas the interviews and observations provided a window into the types of critical assessment-related practices in which teachers engaged, the survey asked only five questions about particular CE component-related concepts.

Surveyed teachers were asked what they would change, if they could change anything about assessment with regard to emerging bilingual students. Their responses revealed that they propose interesting and often evidence-based changes. For example, using native language assessment was a common response. The literature in the field supports this notion, so long as students are literate in their native language (Bowles & Stansfield, 2008). It is also important that students are learning about the tested content knowledge within the context of a bilingual program (Escamilla, personal communication, October 14, 2013). Providing native-language assessment could be a misguided notion if students are receiving all of their instruction in English, as students would not yet have developed the relevant academic language in the content area to support their understanding. Unfortunately, the majority of emerging bilingual students are not exposed to bilingual programs and not all emerging bilingual students develop literacy skills in their first language. This might explain why a relatively higher proportion of CLD-
Educated teachers proposed changes other than native-language assessment, since other changes may be more widely applicable to all emerging bilingual students.

Another commonly proposed change was to make language used on content-area tests more accessible to emerging bilingual students. This is a proposal that could be widely applicable to all students learning English as a new language. Linguistic modification, as this proposal is referred to in the literature, is one of the most promising accommodations in the field of testing accommodations, though it can be difficult to operationalize (Rivera & Collum, 2006). It can also be difficult to justify politically, as people may misinterpret linguistic modification and confuse it with simplifying the content of the test itself. This shows that teachers have some promising ideas about how to improve assessment methods for emerging bilingual students.

The third most common response regarding how to change assessment for emerging bilingual students was to reduce the number of assessments given to emerging bilinguals or to exclude them from some mandated testing. This parallels another response, which was to reduce the importance of testing for emerging bilingual students. These suggestions are apt, given that even the Secretary of Education, Arne Duncan, recently admitted that “testing issues today are sucking the oxygen out of the room in a lot of schools” (Duncan, 2014). He referred to mainstream classes in his discussion, and it is important to remember that mainstream students do not take the additional language-proficiency assessments to which students identified as ELLs are exposed. Teachers’ statements that call for reducing both the time spent testing and its relative importance are not unfounded. One teacher even described how she engaged in a CE component activity after she thought critically about the time she spends administering the “Teaching Strategies GOLD™” assessments: “I have communicated with the GOLD company about my concerns” (Surveyed CLD-teacher 21, personal communication, September 17, 2014).
She explained that the assessments have redundant test items and that they are excessively time consuming for the relative amount of useful information teachers and parents receive from administering the items. She added that she took action on these concerns by communicating with the company itself. This is an exemplification of how teachers can act critically to incite positive changes for emerging bilingual students. She used her agency to voice her concerns over the assessments to the test developers, and hopefully her voice will be heard.

Another CE component-related item on the survey sought to determine if teachers regularly create their own assessments, which could represent a sign of critical activity with regard to assessment for emerging bilingual students. When survey participants were asked if they create their own assessments, most (73%) responded that they create their own assessments “usually,” “frequently,” or “always.” When they do not create their own assessments, the majority of teachers (60%) modify their assessments “usually,” “frequently,” or “always.” However, there is little context provided in teachers’ responses that can explain why teachers create or modify their assessments. Ideally, a teacher who is using assessment critically for emerging bilingual students would create or modify an assessment based on their students’ needs, whether their needs are linguistic, academic, cultural, or socio-emotional in nature.

Teachers were asked to provide an open-ended response to the question, “How do you change or modify assessments?” after they were asked how frequently they modify their assessments. These responses were coded according to the type of modification teachers mentioned in their response. These modification types, along with their relative frequency, are depicted in Table 16. Teachers most frequently reported that they made language-related modifications to their assessments, followed by modifications to test structure and modifications based on students’ needs. Of the teachers who mentioned that they make language-related
modifications, teachers mentioned scaffolding and language supports to lower the language load in the items of their formal assessments. These supports were frequently focused at the word level, such as offering students word banks, changing potentially challenging vocabulary to words with which students are likely to be more familiar, and adding definitions of words. Teachers also mentioned some modifications that target language at the sentence and discourse level, such as adding sentence stems, re-wording items, and focusing more on oral fluency on assessments. Sometimes, teachers’ responses were too vague to interpret how they might make a modification and were not coded. Regardless, the results of the codes developed for the survey question, “How do you modify assessments?” reveal that teachers do account for issues of language and accessibility when modifying their assessments for their emerging bilingual students.

Table 16

*Types of Modifications That Teachers Make to Their Assessments for Emerging Bilingual Students and Relative Frequency of These Modifications*

<table>
<thead>
<tr>
<th>Type of modification</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>language-related</td>
<td>20</td>
</tr>
<tr>
<td>test structure-related</td>
<td>10</td>
</tr>
<tr>
<td>student needs-related</td>
<td>5</td>
</tr>
<tr>
<td>content objective-related</td>
<td>4</td>
</tr>
<tr>
<td>semiotic-related</td>
<td>4</td>
</tr>
<tr>
<td>administration-related</td>
<td>1</td>
</tr>
<tr>
<td>culture-related</td>
<td>1</td>
</tr>
<tr>
<td>authenticity-related</td>
<td>1</td>
</tr>
<tr>
<td>metacognitive-related</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>46</strong></td>
</tr>
</tbody>
</table>

Also significant is the fact that teachers address their students’ needs (n=5) when modifying assessments, which means that teachers are using their System I skills to notice their students’ performance. Then, they use System II skills to interpret their students’ performance
and decide how to respond to push their students further in their zone of proximal development. Finally, they use their CE component skills by modifying their assessments to respond accordingly to their students’ needs. This shows that teachers engage in formative assessment practice when they are assessing their emerging bilingual students.

Formative Assessment as a Process Simultaneously Involving Systems I, II, and the CE Component

In fact, the most common occurrence of CE component-related activities among all teachers involved adapting or transforming instruction and assessment practices across all interviews and observations (though Claire was also reported to have an equal percentage of ‘creating’ as a CE component activity in the interview). Most of these adapting and transforming activities can be attributed to the fact that teachers respond to what they observe about students’ understanding of the content and language objectives. Teachers respond by adjusting their instruction to meet the needs of their students, either through their actions, questions, or materials in class. At times, this adaptation can be quite basic, since almost all teachers shape the feedback they give to students based on what students say or write. However, sometimes this adaptation can be more extensive and complex. Most likely, teachers learn how to increase the complexity of their adaptations or broaden the focus of their adaptations (to include issues of language, for example) with time and experience.

To illustrate this type of event, consider that Alma would occasionally use Spanish with a new emerging bilingual in her class to support her in better understanding a math lesson. She would provide hints to get her student started when it became clear to Alma from her student’s facial expression that she did not know how to begin. She would explain, “Primero, tú haces lo que está al dentro de los paréntesis, y luego haces lo que está fuera. (First, you do what is inside the parentheses, and then you do what is outside)” (Alma, field notes, April 18, 2014). To help
her students access the lesson, she would also encourage her emerging bilinguals to make linguistic connections by explicitly teaching cognates: “In Spanish, do you know what it is called?” she asked. When students gave a confused look, she continued, “paralelogramo (parallelogram)” (Alma, field notes, April 18, 2014). In this math lesson, Alma made use of her students’ facial expressions, and perhaps even their body language, to shift her instructional approach. When working in small groups with her emerging bilingual students, she was able to notice when a student was struggling to complete the class work. She quickly analyzed how to help her student access the lesson better, and adapted her instruction by using native language to supplement the content of her original mini-lesson.

The description of this event in Alma’s class portrays an instance of a formative assessment cycle that addresses elements of language as well. This cycle can be described using component-related activities in ‘becoming familiar’ (a System I activity) with what students understand, ‘analyzing’ (a System II activity) how to respond to a student’s situation if they are not able to entirely access the lesson, and ‘adapting’ (a CE component activity) one’s instructional approach based on this analyzed information. All components of the formative assessment cycle, from collecting information about students’ understanding to acting on that information, must be addressed in order for an instructional event to be considered a good example of a cycle of formative assessment practice. However, when all components of the formative assessment cycle are evident in an instructional event, the event can be applauded as exceptional instructional practice according to assessment experts (Black & Wiliam, 2009; Dunn & Mulvenon, 2009; Filsecker & Kerres, 2012; Ruiz-Primo, Furtak, Ayala, Yin, & Shavelson, 2010).
Conclusion

After examining the results to determine the extent to which teachers of emerging bilingual students use assessment critically, it became apparent that some teachers engage in critical use of assessment to a larger degree than others. Half of the focus teachers stood out as using CE component activities more than the other teachers. Valera, Serena, and Alma all employed a larger percentage of CE component activities, which lead one to conclude that they use assessment more critically.

There are several characteristics that these teachers share that may help to explain how they emerged as more critical users of assessment. First, they all engaged in CE component activities authentically, as they employed critical activities despite the fact that they were not necessarily encouraged or supported by their school culture or administration. Some teachers engaged in CE component activities, but their activities were mandated by their schools. It was challenging to differentiate if their critical use of assessment was derived from their own initiative or if they were simply complying with the norms and values of their schools’ rules and culture. A second commonality was that the critical users of assessment all focused on developing critical thinking skills in their students. These activities served to empower their students, helping them learn to resist oppressive influences they might face in society at large. A third commonality was that all three of the focus teachers who engaged in a higher percentage of CE component activities shared CLD educational experience. While not all of them had finished the program yet, they were all exposed to the CLD curriculum to some extent.

The analysis of the survey items revealed that teachers have some evidence-based ideas about how to change assessment practices used with emerging bilingual learners. The most common responses were to include native language assessments, to make language more
accessible on assessments, and to reduce the amount of testing or exclude emerging bilingual students from some of the testing that takes place. Another survey item demonstrated that most respondents usually create their own assessments. When respondents do not create their own assessments, most usually modify the assessments to a certain extent. This is a positive sign that teachers may use assessment more critically to serve the needs of their emerging bilingual students. Survey respondents also claimed to modify their assessments with issues of language and accessibility in mind. When respondents explained how they modify assessments for emerging bilingual students, many of them mentioned language-related factors to increase the accessibility of the assessment for their students.

Additionally, it became apparent that someone proficient in using assessment critically for emerging bilingual students is also likely to regularly engage in the process of formative assessment. Formative assessment incorporates activities that begin in System I, transfer to System II, and end in the CE component. Teachers who want to become more critical users of assessment for emerging bilingual students might consider beginning by developing habits of formative assessment focused on the area of language.

**Research Question 3: Which Factors Influence Teachers’ Critical Thinking About and Critical Use of Assessment for Emerging Bilingual Students?**

*Focus Teachers’ System I Activities and Related Factors Influencing Their Critical Use of Assessment*

The focus teacher who reported the highest percentage of System I activities in the interview and observation data engages more in thinking that is relatively automatic rather than critical. This automatic kind of thinking is more easily influenced by other factors, though all teachers are exposed to factors that influence their thinking. Based on the interview data, Lucas
was reported to have the highest percentage of System I activities, followed by Claire and Adele. Based on the observation data, Lucas was also reported to have the highest percentage of System I activities, followed by Adele and Claire. With respect to the combined interview and observation data, Lucas again emerged as having the highest percentage of CE component activities, followed by Adele and Claire. Among the interview and observation data sources, the same three teachers were consistently reported to have the highest level of System I activity.

It is interesting that these teachers had fewer years of teaching experience, on average, than the others in the study. Additionally, Lucas, Claire, and Adele all did not have the experience of attending an educational program to become certified in CLD education. Claire was certified by the state of Colorado in CLD education, yet she was not exposed to an educational program for this certification. To merit this certification, she sat for an exam and received a passing score. The remaining three focus teachers all shared CLD program experience. This suggests two potential variables that are positively linked to greater relative use of CE component-related activities: teaching experience and experience in a CLD educational program. Correspondingly, this implies that teachers would engage in proportionally fewer System I related activities. However, it is impossible to differentiate from the results which of these two factors, CLD experience or years of teaching experience, had more influence on this finding.

The types of System I related activities most frequently recorded by practically all teachers in the interviews include ‘complying.’ The only teacher who did not have ‘complying’ as her most commonly recorded System I activity was Alma, and ‘complying’ was still her second most commonly recorded activity. Alma most frequently engaged in the activities of ‘using’ and ‘applying’ instead.
With regard to the observation data, practically all teachers had ‘listening’ recorded as their most common System I activity, with Serena as the only teacher whose most commonly recorded System I activity was ‘becoming familiar’ instead. ‘Listening’ was the second most frequently recorded activity for Serena as well. Therefore, there is a strong trend that teachers mostly discussed activities related to compliance when they were interviewed and that they mostly engaged in activities related to listening when they were observed. This is understandable, given that in an interview, one can discuss the activities in which they take part and why. For many teachers, often the assessment-related activities in which they participated were initiated out of external compliance. In an observation; however, one cannot explain the reasons why one engages in an activity. The observer simply sees the activity taking place. In this case, the most common System I activity involved a teacher listening to her/his students over the course of a lesson. This leads one to conclude that teachers’ critical thinking and use of assessment is most often influenced by either external forces with which teachers are complying or by the needs of students in teachers’ classrooms. It is a positive attribute for teachers to listen to their students in order to inform their critical thinking and action in their assessment process. This is a finding that shows the need for recognition of the work in which teachers routinely engage.

Acts of Compliance

The finding that teachers often act in order to comply with external factors warrants further investigation. In the interviews, teachers described many external forces that caused them to act out of compliance. The assessment- and language-related regulations at schools varied. Several external forces mentioned by teachers were imposed at a school-wide level. Since Valera
and Lucas taught at the same school, as well as Alma and Adele, these external forces were shared for both sets of teachers.

Valera and Lucas both worked in a school in which the grades on the school report cards must match the scores from the high-stakes tests administered in the spring, for example. They complied with this policy each time they wrote their report cards. They also complied each time they recorded a grade in their gradebook based on their own assessment system, knowing it would not be counted towards the final grade on the student’s report card if it did not correspond to the student’s standardized test score. Serena’s school had this same policy, in which the grade on a student’s report card is only reflective of their test scores. In Serena’s school; however, the test scores on the report cards were from school-wide tests given every six weeks instead of state-wide tests given only in the spring. Still, Serena needed to discard grades based on her own assessment practice in order to comply with this policy.

Alma and Adele both worked in a school in which they attended data team meetings that required teachers to administer specific assessments and analyze the data from them in a prescribed manner. All of the six focus teachers had data team meetings, since this was a district-wide initiative. However, the data team meetings were new in Adele and Alma’s school and were being implemented in a relatively more rigorous fashion than in the other schools. In some instances, Alma and Adele found themselves assessing their students in a formal way to prepare for the data team meetings when they would not have otherwise engaged in a formal assessment process. I observed Adele pre-assessing her students’ knowledge of poetry in preparation for a poetry unit she was beginning (Adele, field notes, May 13, 2014). In order to do this, she asked students to write a poem in five minutes and she provided no other scaffolding or instructions for
students. She analyzed their writing with a poetry rubric. Adele would most likely not have carried out this pre-assessment task without the data team meeting requirements.

While the act of pre-assessing is a sign of high-quality practice in theory, the context of assessing a student’s ability to write a poem on demand is rather meaningless. Poets do not write poems on demand (unless it is improvised, which is a sub-category of the genre), nor do authors write their tomes under the pressure of a ticking clock. Assessing students’ writing on demand is meaningful in the context of responding to a content-based prompt or answering questions about a reading passage. This approach reflects the types of writing in which students engage during high-stakes testing situations. However, one must exercise caution in extending the practice of administering on-demand writing assessments to more creative pieces of poetry or narratives. One cannot evaluate the quality of a poem based on a strict numeric rubric.

Data-Driven Focus Limits Teachers’ Initiatives for Critical and Creative Pathways

Alma expressed frustration in complying with the idea that students need to be taught writing in this fashion, since she recognized that quality creative writing does not emerge from this type of “numbers game.” In addition, she observed that students engage in fewer creative writing tasks since the practice of data-driven instruction became a school-wide initiative. Students also have fewer opportunities to engage in social justice-oriented topics during English Language Arts due to the data-driven initiatives. The compliance with the data-driven instruction and assessment policy imposed limitations for Alma in terms of her ability to find time to engage in writing activities that ask students to think more critically. That said, Alma and Adele both saw value in the data-driven initiative, since it would help address the shift in priorities to achieving high scores on standardized tests. Alma explained how, “with this shift…. it’s been more of, I think they’re what, it’s definitely thinking with the end in mind, standardized tests,
what is that going to look like, and how are we going to support them to pass that” (Alma, personal communication, April 18, 2014).

In the context of a standardized testing era, little emphasis is placed on students being able to cultivate skills such as creativity. When poetry and narratives are graded under the guidance of strict rubrics, there is little opportunity for students to think “outside the box” to stretch their imaginations. For instance, Alma argued: “[W]ith the writing piece, we’re looking at, we’re really prompting them. We use prompts to gather data and I feel like there’s been this shift between, you know, really building their stamina and also building their creative capacity” (Alma, personal communication, April 18, 2014).

Alma expressed that assessing students’ writing becomes a numerical process, in which they must assess by identifying which parts of the writing rubric students are struggling with. This practice is required school-wide because “that’s the rubric used for standardized testing” (Alma, personal communication, April 18, 2014). It is important for students to learn this approach to writing. Yet, it is equally imperative that students be given opportunities to build their creative capacity. When a teacher tells students, “‘OK now you’re gonna spend time writing’ and they look at you and they say, ‘Well what’s the prompt?’” it is clear that students are mainly learning to respond to prompts instead of being given opportunities to experiment with writing for a variety of purposes (Alma, personal communication, April 18, 2014). Experimentation with language is particularly important for students learning English. Sir Ken Robinson agrees that this is a problem. His 2006 TED Talk on the importance of valuing creativity in schools is the most watched TED Talk presentation in history, with more than 32 million views (CBS, 2015). He claims that “most teachers I know don’t like this either, but there
is this deadly culture of standardizing that’s being pushed on them from, um, politically. And my core message here is that we have to personalize education, not standardize it” (CBS, 2015).

Standardized testing is a federal- and state-imposed, mandated practice. In the interviews, focus teachers often referred to compliance with the external forces of standardized testing-related initiatives as well. For example, the focus on standardized testing at the national level has trickled down to testing-related initiatives at the district level in many areas. Since funding for schools is dependent, to a certain extent, on the performance of a school’s students on standardized tests, the priority of many schools revolves around improving students’ test scores. Schools are threatened to be closed down by the state if they do not demonstrate sufficient progress and performance on standardized tests as well. The priority of standardized testing performance has therefore affected both the content that is taught and the manner in which content is taught.

**Data-Driven Focus Does Not Prioritize Students’ Socio-Emotional Development**

It is crucial to consider the needs of students in socio-emotional areas as well as in the academic realm. Alma described the significance of fostering students’ socio-emotional development when she recounted how she allows a few additional minutes in the morning for students to unpack so that she can check in with them to see, “OK, who looks like they’ve had a rough night?” because “they have to be learning ready” (Alma, personal communication, April 18, 2014). This is particularly important given the experiences of many emerging bilingual students, who need to lower their affective filter in order to make progress in their language development in the school environment. These same students are, unfortunately, often the ones who must overcome prejudice and low expectations from their teachers, their peers, and society due to their language status. Therefore, it is of utmost importance for teachers to critically assess
socio-emotional aspects of learning for emerging bilingual students as well, as Alma does in her practice.

*Curriculum Derived From Testing as a Limiting Factor for Critical Thinking and Action*

The content that is taught by the focus teachers reflects primarily what is assessed by standardized tests. This can be positive when the assessed standards cover everything one would want students to understand over the course of the school year. However, since some content is not easily assessed by large-scale testing systems, teachers sometimes feel it is their duty to find slivers of time between teaching the mandated standards in order to teach important concepts that are not assessed. All students, but particularly language learners, benefit when teachers bring in and validate their background knowledge about a topic and then link this knowledge to the standard that is to be taught. For example, language learners can benefit from learning about topics that are taught with culturally relevant context.

Even broad content areas, such as science and social studies, are relegated to the background if they are not tested. During the NCLB era, only literacy and mathematics was assessed on large-scale tests. As a result, primarily literacy and mathematics have been taught in some schools, particularly those schools that struggle to demonstrate adequate performance on standardized tests. Lucas explained that,

[I]t’s always been reading, writing, math, and it’s never been science and social studies in there. But now science and social studies is gonna be tested and it’s gonna be accounted for, so we’re gonna have to make time for that. Big time. And like we’ve talked with all the other teachers and that’s the same case with all of ‘em. It gets pushed aside from Kindergarten up, so when it comes to 5th grade and they don’t know the water cycle, it’s kinda tough. (Lucas, personal communication, April 30, 2014)
Now that large-scale testing systems are beginning to incorporate the content areas of science and social studies, there is a shift to focus more on these content areas. Teachers are finding this to be challenging; however, given that many students have never been taught in depth about these subjects before. If teachers were given opportunities to teach and assess critically without great pressure from testing, they would not be likely to impose limits on covering any content area. A strict emphasis on large-scale testing imposes limits on teachers in terms of content objectives, let alone language objectives, which can prevent them from having the freedom to think and act critically with regard to their instruction and assessment for emerging bilingual students.

_Data-Driven Focus Limits Teachers’ Critical Curriculum_

Often there is a lot of content to cover for the large-scale testing sessions, with little room for other issues like culturally responsive instruction or social justice initiatives. Alma exemplified this phenomenon in her classroom when she described “carving out time for a read aloud and really talk[ing] in depth about those social justice kind of things so that education is still transformative in nature” (Alma, personal communication, April 18, 2014). Alma planned to select topics in her poetry unit that were relevant to her students’ lives also. For instance, she expressed interest in showing the film, “Cesar Chavez,” and following that with another recent movie, “The Watsons go to Birmingham”. Both of these topics touch on social justice issues relevant to her students’ lives, in addition to the fact that they are newly released movies in popular culture. Teaching social justice or other critically minded topics are beneficial in helping students to become more empowered as learners and as citizens in general.

Teaching these “extras” is not necessarily encouraged by the administration of the focus teachers, as teachers find themselves needing to choose between omitting the teaching of certain
critical concepts or complying with policies. For example, Serena shared that there is “definitely a tension between what I want my classroom to look like, and how they are, and what their grades look like now” (Serena, personal communication, May 16, 2014). She explained that their grade was solely based on the tests given every six weeks, and the tests are modeled after the state standardized testing program. Due to this, the projects that promote social justice initiatives or that focus more on language cannot count towards students’ grades for the class. Since these types of initiatives do not contribute to the evaluation of educators’ teaching and students’ learning, there is no external benefit for Serena or her students to focus on them. In fact, when Serena spends time on social justice-oriented projects with her students, she risks their potentially lower performance on the specific content objectives on the unit tests.

Data-Driven Focus Encourages Learning Outcomes at the Expense of Learning Processes

The priority of standardized testing affects the manner in which content is taught to students. The scoring of large-scale, standardized tests is performed in a manner in which the process can be efficiently and quickly completed. To facilitate this, items on these tests are frequently in closed-response formats (e.g. multiple choice, cloze, etc.) or in open-response formats graded by a rubric. They are scored with a strict rubric in order to achieve high reliability across raters of the test, as reliability of such a large-scale test is important. However, smaller scale tests and even classroom-based assessments are now modeled after these large-scale assessments, with a priority on closed-response formats and open-response formats graded by a rubric. While it is important for students to learn how to address specific content, applying this approach to all forms of assessment devalues the process of learning. Serena resists this approach, explaining that on her assessments, students “get points for doing these different things, because I want to show them it's about the process, it's not all about the right answer”
(Serena, personal communication, May 16, 2014). However, a focus on the process of mathematics is not incentivized when the test, which consists of mostly closed-response items, makes up the students’ entire grade for the class.

When students’ standardized or unit test score comprises their only grade for the class, as is the case in all focus teachers’ classrooms except for Adele and Alma, it communicates a strong, implicit message that teachers’ assessments simply do not matter. The informal assessments conducted by teachers often serve to focus on the process of learning. This can include a teacher’s informal diagnostic assessment of content vocabulary when students begin a unit, or it can involve a teacher taking note of collaborative classroom conversations. Serena shared how, “every conversation, all of the work that they're doing is basically an assessment for me of how, ‘Where are their misconceptions?’ Where do we need to help them, and moving forward, they're able to create a product that they actually care about” (Serena, personal communication, May 16, 2014).

However, due to school policies, these informal, formative assessments do not contribute to the grade a student receives. Ultimately, if a teacher’s assessment methods are not valued enough to form a fraction of their students’ grades, then they will learn to place little value in the more informal, formative assessments of their students’ progress. These are precisely the type of assessments that have potential to help students improve their skills. Standardized testing and similarly-structured assessments mostly serve an evaluative purpose. Helping students to learn is not a primary goal of this type of testing. It is hard for teachers to determine how to help students based on these types of assessments, as teachers are only able to know whether or not students got the right answer instead of being able to pinpoint the step in the process where they faltered.
When a heavy emphasis is placed on standardized testing structures in assessment, the process of learning becomes lost in the shuffle of scantrons.

System I Survey Findings on the Relative Value of Classroom and Large-Scale Assessments

Given that students’ grades for three out of six of the focus teachers are comprised solely from large-scale, standardized tests or end-of-unit tests, it is evident that focus teachers’ schools are placing value mainly on summative assessments. However, the focus teachers mostly communicated that they emphasize formative assessments in the process of students’ learning, even when their schools do not encourage this type of assessment.

To determine if the survey results reflected this trend, I analyzed the first two System I related survey items. The first item revealed that only 13% of surveyed teachers felt that summative assessments, in the form of classroom tests, provided them the most helpful information regarding their emerging bilingual students’ progress towards content area learning goals. The remainder of teachers selected that formative assessments (in the form of quizzes, in-class activities, and conversation in class) provide them with the most helpful information. No respondents selected that large-scale tests give them the most helpful information in this regard. The second item revealed that only 2% of surveyed teachers felt that summative assessments, in the form of large-scale tests, provided them the most helpful information regarding their emerging bilingual students’ progress towards language proficiency goals. The remainder of teachers selected that formative assessments (in the form of homework, quizzes, in-class activities, and conversation in class) provide them with the most helpful information.

This shows that teachers mostly value informal formative assessments when they are attempting to ascertain where their students are in their progress towards their content area- and language-learning goals. When teachers were asked if their administration thinks classroom
assessments are important, however, only 50% of them responded that their administration considers classroom assessments to be ‘always’ or ‘frequently’ important, whereas 67% of teachers thought their administration considers large-scale assessments to be ‘always’ or ‘frequently’ important. Likewise, when teachers were asked if their district thinks classroom assessments are important, only 47% responded that their district considers classroom-based assessments to be ‘always’ or ‘frequently’ important, whereas 82% thought their district considers large-scale assessments to be ‘always’ or ‘frequently’ important. The survey results show that teachers perceive that their administration and their district do not prioritize classroom assessments nearly as much as they do for large-scale assessments. At the same time, 78% of teachers responded that they think classroom assessments are ‘always’ or ‘frequently’ important. Only 18% of surveyed teachers responded that they think large-scale assessments are ‘always’ or ‘frequently’ important. The survey results support the assertions made by the focus teachers that classroom-based assessments are more valued by teachers than by their schools or districts.

Schools and districts tend to assign stronger priority to large-scale assessments instead.  

*Distribution of Large Scale Assessment Results as a Limiting Factor*

Although some classroom assessments are summative in nature, most classroom assessments are formative because the teacher can use their students’ assessment information to inform their subsequent instruction. Even after a unit is complete, teachers might “spiral” some of the skills that were not mastered on a unit test into the following units of study. Large scale tests are typically administered near the end of the school year, with the results being returned once the students have completed that grade level. Teachers, by default, cannot use the results of many large-scale tests to inform their instruction for their students. They might be able to inform their instruction with students for the following year in general, though they are not able to
personalize instructional changes for their current students using large-scale test results. Teachers find it useful to have classroom assessment results immediately, so they are able to tailor their instruction for their students based on that information. Interim tests do provide some timely information, though the purpose of interim testing is still primarily evaluative rather than instructive.

A significant factor that prevents teachers from being able to use assessment critically for their emerging bilinguals is the lack of usability and relevance of large-scale testing results once they are distributed. In fact, Valera discussed how the last benchmark test is “summative, and then it’s the end of the year” (Valera, personal communication, April 14, 2014). She insisted that, with regard to all large-scale tests, “if I had data or if I was able to use the data from those tests, then I would put more stock into looking at the outcomes” (Valera, personal communication, April 14, 2014). She did acknowledge that she uses the first two benchmark tests administered in fall and spring, since the timing allows for her to use the information.

Survey Findings on What Prevents Teachers From Making Changes

When surveyed teachers were asked what holds them back from making positive changes for their emerging bilingual students, they listed many factors that impinged on their ability to act. The most common coded response was that school, district, or state requirements hold them back from making changes. Figure 26 shows the responses from all surveyed teachers’ coded responses.
Figure 26. Surveyed teachers’ coded responses to the question: “What holds you back from making changes?” percentages.

Twenty-eight percent of all surveyed teachers’ coded responses fit this response type, and 14 teachers specifically referred to large-scale testing requirements in their responses. One teacher who recommended providing testing accommodations tailored to students’ language levels wrote, “In my classroom, I already make these changes - we have not been able to do that on high-stakes tests due to regulations” (Surveyed non-CLD teacher 11, personal communication, October 16, 2014). This teacher uses her critical thinking about language in assessment to make changes in her own classroom, but is not allowed to make these changes in large-scale assessment scenarios. Twenty percent of teachers’ coded responses specifically described having a lack of power or control over assessment for emerging bilingual students. Consider that the surveyed teachers who listed ‘requirements’ as a barrier to making change also likely share feelings of powerlessness due to those requirements. If this is the case, then approximately half of all surveyed teachers’ coded responses expressed a lack of power or control in some form.
While there will always be regulations regarding assessment, the regulations should be perceived as fair for most of the teachers who must follow them. However, the survey results paint a different picture. Many teachers voiced concerns about large-scale testing, yet they expressed feelings of powerlessness about these kinds of required assessments. Teachers’ comments can be summarized in the word graphic in Figure 27, which was created using the web-based tool, “Word it out”. The word graphic is presented to illustrate teachers’ responses describing feelings of disempowerment with regard to assessment for emerging bilingual students. If teachers feel this powerless, one can only imagine how disempowered emerging bilingual students must feel about the assessment process in their schools. This is an issue that must be addressed if schools are to act as structures to engage and free students from subordination. Teachers and students must be comfortable to think critically and act with agency. If educational leaders do not address this, schools will continue to reproduce the inequities that exist in society at large.

Figure 27. Word graphic displaying surveyed teachers’ comments regarding their lack of power or control with assessment for emerging bilingual students.
The other factors that teachers listed when asked what holds them back from making changes included legal or political issues. Twelve percent of teachers’ coded responses mentioned legal or political barriers to making changes, which also represents a lack of control or power. Teachers who voiced legal or political barriers did not explicitly relate these concepts; however, testing requirements often stem from educational legislation. Thus, teachers who responded in this manner are also experiencing similar feelings regarding assessment for emerging bilingual students.

Internal Factors Precluding Teachers from Critical Thinking and Action

One internal factor that possibly prevents teachers from thinking and acting critically is that of time. Seventeen percent of surveyed teachers’ coded responses expressed time as the culprit in holding them back from making changes in assessment for emerging bilingual students. Teachers communicated various types of productive endeavors in which they would engage if they had the time. Some of these endeavors include: “time to analyze the questions to see what might trip them up,” “include cultural and relevant tasks,” “having time to plan language assessments,” “use more visuals,” and “alignment [of large scale ELL assessments] with classroom instruction.” All of these activities take time, which must come after hours and in addition to what the content area teacher has planned for her/his instruction for the next day. If teachers were given some time during their work day to collaborate or work individually on assessment for emerging bilingual students, they could quickly become more critical users of assessment.

Another internal factor that could preclude teachers from engaging in critical use of assessment with their students is the fear that their students may learn to be critical of them. Valera voiced the idea that teachers might be afraid of modeling critical use of assessment, and
empowering their students to think and act critically, since students might point out inconsistencies about their teaching. It can be intimidating for teachers, who are in a position of power over their students, to promote critical thinking because they are inviting critique of themselves. The same can be said of school administrators in encouraging critical thinking in their teachers, or of district and state officials in endorsing critical thinking in their schools.

There is an implicit transition of power from a hierarchical to a more equitable structure when critical thinking and action is encouraged, since those who are learning to think and act more critically gain agency. This transition can be threatening to those in power, and it is human nature to resist this threat. However, this transition of power helps lay the groundwork for more equitable schools, and eventually, a more equitable society.

CLD Educational Experience as a Factor in Teachers’ Critical Use of Assessment

With regard to the focus teachers that were interviewed and observed, the results support the claim that CLD educational experience is helpful in developing teachers’ critical use of assessment. The three focus teachers with CLD educational experience were Alma, Serena, and Valera. The three focus teachers without this experience were Adele, Lucas, and Claire. The CLD-educated teachers were reported to have a higher mean percentage of CE component activities and a lower mean percentage of System I activities than the teachers without CLD educational experience. Figure 28 represents this data visually.

Interestingly, both groups of teachers were reported to engage in similar percentages of System II activities. This was an unexpected result, as one might anticipate that with additional educational experience in CLD education, one might think more critically about issues of assessment with emerging bilingual students. Though there is a small positive difference in System II thinking between CLD-educated teachers and those without this training, one may
wonder if changes to CLD programs could help CLD-educated teachers increase their System II thinking. It is possible that the measurement tools are not refined enough yet to capture the differences among teachers, or perhaps the professional development trainings that non-CLD-experienced teachers receive help them to develop critical reasoning skills about aspects of assessment for emerging bilinguals. Regardless, most all teachers stand to benefit from further development of their critical thinking skills in this area.

There is evidence to suggest that, while teachers with CLD educational experience do not think significantly more critically than teachers without this experience, CLD-educated teachers do engage in critical activity more frequently. This corresponds to a decrease in System I activities for CLD-educated teachers and an increase in System I activities for teachers without this experience.

![Figure 28](image.png)

**Figure 28.** CLD and Non-CLD focus teachers’ relative engagement in System I, System II, and CE component activities from the combined interview and observation data in percentages.

The specific types of activities each teacher used within each component are shown in Table 17. One may notice that the teachers with CLD educational experience have a lower mean percentage of compliance-related activities, whereas teachers without this experience spend
relatively more time engaging in compliance-oriented activities. The CLD-educated group of teachers also has a lower mean percentage of the System I activity, ‘assuming.’ This shows that CLD-educated teachers may spend a little more time developing their own opinions about a concept, rather than making assumptions about it or simply complying with it.

With regard to CE component differences among teachers, CLD-educated teachers spent relatively more time involved in the activities of “criticizing” and “questioning” than non-CLD educated teachers. They also had a higher mean percentage of the CE component activity, “creating.” However, they did have a slightly lower mean percentage in the activities of “adapting” and “transforming” as a proportion of their total CE component activities. This was unexpected, but it is a slight difference and, still, the absolute magnitude of CLD-educated teachers’ CE component activities exceeds the non-CLD-educated teachers’ CE component activities. It is also possible that adapting one’s instruction to the general needs of emerging bilingual students is a somewhat common classroom practice, since most teachers modify their instruction based on how students respond to their questioning. Going forward, a more strict definition of “adapting” and “transforming” may need to be devised. That would allow an observer to discern which activities of adaptation and transformation were specifically targeted at linguistic aspects for emerging bilingual students.

Finally, the mean percentage attributed to the activity of “collaborating” by CLD-educated teachers was 16% lower than that of non-CLD teachers. It is surprising that CLD-educated teachers would report that they collaborate less frequently than non-CLD-educated teachers. However, in the interviews, non-CLD experienced teachers Adele and Lucas both expressed a desire to collaborate with other teachers for support, since they were newer to the profession. This could have had a rather large impact on the mean percentages reported for
teachers’ collaboration levels. Ideally for the purposes of analysis, the CLD-educated teachers would have been paired with non-CLD-educated teachers with similar years of teaching experience. This would have limited differences due to varying levels of teaching experience, which serves as a confounding factor in this case.

Table 17

*Component-Related Activities of Focus Teachers Grouped by CLD Educational Experience in Percentages*

<table>
<thead>
<tr>
<th>System I</th>
<th>Mean CLD Teachers</th>
<th>Mean Non-CLD Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Become familiar</td>
<td>8.11</td>
<td>6.73</td>
</tr>
<tr>
<td>Listen</td>
<td>5.08</td>
<td>1.86</td>
</tr>
<tr>
<td>Become familiar and listen</td>
<td>4.17</td>
<td>4.79</td>
</tr>
<tr>
<td>Use</td>
<td>3.57</td>
<td>3.72</td>
</tr>
<tr>
<td>Apply</td>
<td>0.00</td>
<td>2.06</td>
</tr>
<tr>
<td>Use and apply</td>
<td>14.24</td>
<td>11.23</td>
</tr>
<tr>
<td>Know</td>
<td>9.66</td>
<td>1.86</td>
</tr>
<tr>
<td>Understand</td>
<td>6.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Know and understand</td>
<td>6.69</td>
<td>17.88</td>
</tr>
<tr>
<td>Comply</td>
<td>38.38</td>
<td>43.24</td>
</tr>
<tr>
<td>Copy</td>
<td>3.03</td>
<td>1.95</td>
</tr>
<tr>
<td>Assume</td>
<td>1.01</td>
<td>4.69</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>System II</th>
<th>Mean CLD Teachers</th>
<th>Mean Non-CLD Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compare</td>
<td>2.63</td>
<td>4.53</td>
</tr>
<tr>
<td>Contrast</td>
<td>9.43</td>
<td>6.58</td>
</tr>
<tr>
<td>Analyze</td>
<td>7.89</td>
<td>10.03</td>
</tr>
<tr>
<td>Interpret</td>
<td>11.58</td>
<td>11.26</td>
</tr>
<tr>
<td>Analyze and interpret</td>
<td>0.00</td>
<td>4.17</td>
</tr>
<tr>
<td>Evaluate</td>
<td>6.81</td>
<td>9.26</td>
</tr>
<tr>
<td>Judge</td>
<td>3.88</td>
<td>7.46</td>
</tr>
<tr>
<td>Integrate</td>
<td>6.16</td>
<td>4.02</td>
</tr>
<tr>
<td>Conclude</td>
<td>16.66</td>
<td>13.38</td>
</tr>
<tr>
<td>Reflect</td>
<td>34.96</td>
<td>29.31</td>
</tr>
<tr>
<td>CE component</td>
<td>Mean CLD Teachers</td>
<td>Mean Non-CLD Teachers</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Propose</td>
<td>4.32</td>
<td>9.05</td>
</tr>
<tr>
<td>Imagine</td>
<td>4.32</td>
<td>2.15</td>
</tr>
<tr>
<td>Propose and imagine</td>
<td>2.05</td>
<td>1.08</td>
</tr>
<tr>
<td>Adapt</td>
<td>6.13</td>
<td>10.20</td>
</tr>
<tr>
<td>Transform</td>
<td>0.76</td>
<td>1.15</td>
</tr>
<tr>
<td>Adapt and transform</td>
<td>25.99</td>
<td>23.62</td>
</tr>
<tr>
<td>Criticize</td>
<td>17.59</td>
<td>9.97</td>
</tr>
<tr>
<td>Question</td>
<td>8.94</td>
<td>2.22</td>
</tr>
<tr>
<td>Criticize and question</td>
<td>0.00</td>
<td>1.08</td>
</tr>
<tr>
<td>Organize</td>
<td>3.54</td>
<td>2.22</td>
</tr>
<tr>
<td>Collaborate</td>
<td>6.40</td>
<td>22.47</td>
</tr>
<tr>
<td>Create</td>
<td>19.95</td>
<td>14.79</td>
</tr>
</tbody>
</table>

The survey results tell a similar story about the differences between CLD-educated teachers and those without this experience. Overall, CLD-educated teachers reported they feel relatively more prepared to design high-quality assessments for emerging bilingual students, as one might expect. The majority (67%) of CLD-educated teachers feel either “very” or “mostly” confident to design assessments for emerging bilingual students, whereas only 43% of non-CLD-educated teachers feel this way. Since CLD-educated teachers stated that they feel more prepared in this area, there is an implicit acknowledgment that they have developed some knowledge and understanding about designing assessment. ‘Knowing’ and ‘understanding’ are both System I related activities. These are valuable System I activities that one might encourage teachers to develop in preparation for teaching and assessing emerging bilingual students.

With regard to System II related survey items, the results showed that CLD-educated teachers often use helpful information about ELLs’ progress in their class to make inferences about what to do for students in future language instruction. They selected this option more frequently (33%) than any other option. Teachers without CLD educational experience mostly selected (47%) that they use helpful information about ELLs’ progress to make inferences about
what to do for students in future *content area* instruction. This shows that CLD-experienced teachers more persistently consider aspects of language in addition to content when they are assessing their students.

Interestingly, when teachers were asked what they thought assessment meant to them, a large majority of CLD-experienced teachers (60%) responded that they perceive assessment to be part of their daily practice. Non-CLD-experienced teachers also selected this option most frequently (43%), but to a lesser degree than the CLD-experienced teachers. A moderate percentage (27%) of non-CLD-experienced teachers selected that they think assessment is a way to learn about their students and what their students can do, but that it’s hard to know how to use that information. This result suggests that CLD-experienced teachers may be more equipped to analyze assessment methods and adapt their future instruction accordingly. Since the CLD education program at the University of Colorado at Boulder has an assessment course, as do most CLD education programs, it is logical that teachers who have taken CLD coursework are better poised to use assessment information for emerging bilingual students.

An examination of the CE component-related items on the survey reveal some unexpected findings. Though CLD-educated teachers stated that they felt relatively more prepared to design and implement high quality assessments for ELLs, they responded that they create their own assessments relatively less frequently than their non-CLD-educated counterparts. Only 30% of CLD-educated teachers selected that they “always” or “frequently” create their own assessments, whereas 67% of non-CLD-educated teachers did. Furthermore, only 30% of CLD-educated teachers claimed they change or modify assessments they do not create themselves, whereas 47% of non-CLD-educated teachers did. It is possible that teachers without CLD education experience are more comfortable creating and modifying formal
assessments to test content knowledge, while CLD-educated teachers feel more comfortable creating and modifying assessments to test language understandings. However, since these two survey items did not distinguish between the various potential purposes of assessment, it is impossible to discover if this explanation is warranted without further investigation. It is also possible that CLD-educated teachers may feel more prepared to design and implement assessments for emerging bilingual students, yet they do not follow through on this by creating their own assessments. If this is the case, it would be important to determine what prevents CLD-educated teachers from creating their own assessments.

Additionally, an unanticipated finding emerged in the survey item that asked teachers if they collaborate with their colleagues to plan their assessment methods. The majority of non-CLD-educated teachers (60%) claimed that they collaborate “always” or “frequently” with their colleagues to plan their assessment methods, whereas only 37% of CLD-educated teachers did. The activity of collaborating is a desirable characteristic of a teacher who uses assessment critically, as collaboration can help even the best critical thinkers to discover new ways of thinking about a topic. It warrants further investigation to determine why CLD-educated teachers opted to collaborate relatively less frequently with their colleagues on matters of assessment.

**Conclusion**

After analyzing the results of the data from the focus teachers, it became evident that half of the focus teachers engaged relatively more frequently in System I related activities. Lucas, Claire, and Adele focused more on System I activities, on average. These three teachers shared some commonalities, including the fact that none of them had CLD educational experience and that they had fewer years of teaching experience on average. Across practically all focus teachers, the most commonly recorded System I activity in the interviews was ‘complying,’
whereas in the observations the most commonly recorded System I activity was ‘listening.’ These differences could be due to the context of the varying data sources, since one might see a teacher listen to her students much more frequently in an observation of her teaching than in an interview with her. In the interviews, teachers spoke more frequently about compliance, since they had the opportunity to explain how they comply with regulations related to assessment and emerging bilingual students. Providing report-card grades that match test scores was a common requirement. Preparing data and attending regular data team meetings was another recurring activity that was completed out of compliance.

The focus on data, as communicated through assessment-related requirements at the school, district, or state level, served to limit teachers’ CE component activities. Teachers found themselves with limited time to engage in activities that foster creative and critical thinking in their students. Teachers also found it difficult to foster critical thinking and attend to students’ socio-emotional development due to compliance with a rigid, data-driven curriculum. Teachers are asked to teach only the content that is assessed on large scale tests, either explicitly by their administration or implicitly by having their evaluation based on their students’ scores on the large scale tests. This can render it challenging for teachers to engage in depth with students in culturally relevant instruction, social justice-oriented concepts, or concepts related to students’ socio-emotional development.

The focus on data-driven initiatives also limited teachers’ engagement in critical activity due to the corresponding emphasis on learning outcomes as opposed to learning processes. The “numbers game,” as Alma labeled this phenomenon, prioritizes student responses that can be easily scored with a rubric or an answer sheet. At times, this takes place at the expense of recognizing students’ relative progress in learning the concept at hand. The “numbers game” can
give teachers the sense that their informal, formative classroom assessments are not valued. Informal, formative assessments naturally emphasize the processes of learning, whereas summative assessments serve a more evaluative purpose.

The focus teachers and the surveyed teachers both suggested that they believe their school leaders position large scale assessments as being more valuable than classroom assessments. There is evidence to suggest that the focus teachers and surveyed teachers both value their classroom assessments over large scale assessments, as they garner the most helpful information from classroom assessments. Teachers portrayed a tension between the views of teachers and school management (administration and district personnel) on the relative value of classroom assessments and large scale assessments. Focus teachers also communicated that the untimely distribution of many large scale assessments is a limiting factor in their use. It is challenging to use assessment critically if one does not have access to the information in a timely manner.

Furthermore, the most common factor that prevents surveyed teachers from making changes in assessment for emerging bilingual students is that of school, district, or state requirements. Lack of power or control was the second most common factor listed, and legal or political issues was third. An implicit lack of power or control is present in all of these common responses, revealing that teachers are feeling a lack of control over assessment issues in their own classrooms. There were also some internal factors that teachers perceive to be preventing them from using assessment critically for emerging bilingual students. Some surveyed teachers listed time as a limiting factor. One focus teacher suggested that some teachers may fear introducing a critical outlook into their classroom, since students may become critical of their actions.
Finally, I analyzed the data to determine the effect of CLD educational experience as a variable in teachers’ critical use of assessment. There were mixed results across various data sources. This analysis provided evidence to support the claim that CLD educational experience had a mostly positive effect on critical use of assessment for emerging bilingual students within the focus teacher group. In the group of surveyed teachers, there was some evidence to support this claim, but also some evidence to support the idea that CLD educational experience did not have a positive effect on critical assessment use. The variable of CLD educational experience was not the primary purpose for this study, so it is not surprising that further investigation is warranted in this regard.
Chapter 6
Discussion

In this chapter, I first discuss the study findings. Next, I provide an evidence-based profile of a critical user of assessment based on the evidence from the study and compare it to the theoretical profile of a critical user of assessment established earlier in this work. Then, I identify the potential implications of this study for policy. Additionally, I describe the limitations of this study and propose ideas for future, related research. Finally, I end with closing remarks regarding what I have learned in the process of conducting this research.

Research Question One: To What Extent do Teachers of Emerging Bilingual Students Think Critically About Assessment?

The findings in this study supported the idea that Kahneman (2010) put forth regarding System I and System II thinking processes. The results from the focus teachers showed that, on average, there were fewer pieces of evidence of cognitive activity related to System II than System I. This was particularly evident in the observational data, as the format provided less opportunity for teachers to share their thought processes. Also, the lower incidence of System II thinking could be attributed to the social desirability of responding as though one thinks more critically by employing System II activities in the interviews. Based on Kahneman’s theory, one might have even hypothesized that there would be a lower proportion of System II activities than was apparent in the results. However, Kahneman was clear in his work about how he designed experiments for the express purpose of “trapping” people into using their System I processes when they should have thought more critically about a certain situation. Since my study did not set out with the intention to creative cognitive traps in the interviews or observations, it is logical that there was a higher incidence of System II thinking present in the data than one might expect after reading about Kahneman’s experiments. Unfortunately, none of the focus teachers appeared
to have significantly more expertise in engaging in System II activities than any other focus teacher. However, the surveyed teachers did claim that they engaged in System II activities, such as using information garnered from assessments to inform their subsequent instruction for their language learners.

I identified several areas in which teachers may benefit from further support in professional development. First, teachers may require additional support to help them learn how to use language-related information from assessments to inform their subsequent instruction. Teachers also might need support to determine how to use information from assessments to guide their next lessons in general. In order to further develop their habits of critical thinking, teachers might also need support in establishing metacognitive awareness about how language can interact with the validity of assessments, as well as how to properly respond to this interaction. Finally, teachers may need to be careful not to fall into the trap of using the representativeness heuristic when they are building their expectations for their language learners. This final recommendation may prove to be the most difficult, since expectations of others can be formed in our subconscious minds. Expectations often reflect philosophical and political ideologies, which can be deeply rooted in our belief system and therefore very slow to change. It is important to mention that these suggestions for professional development can and should be implemented by teachers who have mastered these skills. An essential element to developing teachers’ critical use of assessment is to empower teachers in their practice. If teachers can lead other teachers in developing tools to think and act critically with regard to language in assessment, then they will be positioned more as independent critical minds responsible for instilling these traits in the future generation than as managers of a set curriculum that they must deposit into students’ heads (K. E. Johnson, 2006).
Research Question Two: To What Extent do Teachers of Emerging Bilingual Students Use Assessment Critically?

The findings from this study reveal that some teachers use assessment critically for emerging bilingual students more than others. Half of the focus teachers demonstrated higher average levels of critical assessment use, which caused me to search for commonalities among these teachers that might be related to their critical use of assessment. These commonalities helped me to build a more comprehensive profile of the characteristics of a critical user of assessment. The profile of a critical user of assessment is described at length further in this chapter. Additionally, surveyed teachers demonstrated that they use assessment critically around issues of language and assessment. This was evident when teachers described evidence-based ideas regarding how to positively change assessment practices for emerging bilingual learners. A majority of surveyed teachers also described that they create their own assessments, or that they modify them if they use externally-developed assessments. Often teachers cited language or accessibility issues as a reason to modify an assessment.

Research Question Three: Which Factors Influence Teachers’ Critical Thinking About and Critical Use of Assessment for Emerging Bilingual Students?

There is evidence to suggest that some of the focus teachers in this study employed System I activities proportionally more than the others, and I searched for commonalities among these teachers to look for factors that might influence these teachers’ tendencies to use more System I activities. It was interesting that the most frequently used System I activity was ‘complying’ in the context of the interviews and ‘listening’ in the context of the observations.

The issues that emerged when teachers spoke about compliance in the interviews provided me with information regarding the factors that influence teachers’ critical use of assessment. The assessment-related requirements that place an emphasis on students’ data often
served to have a limiting influence on teachers’ critical use of assessment for an array of reasons. Surveyed teachers listed many requirements that prevented them from making positive changes in their assessment use for emerging bilingual students. With a focus on standardized testing data, teachers expressed that they had less time to highlight social justice issues. They communicated that they had limited time to foster critical and creative thinking skill development for their students. Teachers also felt that there was too little time allotted in the curriculum to be able to allow for culturally relevant instruction or to attend to students’ socio-emotional development properly. Teachers also expressed that, with the focus on standardized testing data, much greater stress is placed on learning outcomes than on learning processes. All of these aspects, from creative thinking to considering students’ socio-emotional development, are important aspects that contribute to students’ learning. Critical users of assessment aim to empower students by teaching all these important aspects, particularly students marginalized by language status, in order to build a more equitable system of education.

The “numbers game” approach to learning, as Alma termed the focus on standardized testing to drive instruction and assessment, serves to dehumanize students by reducing them quite literally from individuals to numbers. Teachers who use assessment critically work to empower their students as independent and capable bilingual learners, and an overemphasis on student learning in terms of numbers fundamentally misses opportunities both to recognize emerging bilinguals’ cultural and linguistic capital as well as to further build on this capital. The “Petit Prince” [Little Prince] (de Saint-Exupery, 2015) provides a good metaphor to represent this phenomenon:

Grown-ups like numbers. When you tell them that you have made a new friend, they never ask you any questions about essential matters. They never say to you, ‘What does
his voice sound like? What games does he love best? Does he collect butterflies?’

Instead, they demand: ‘How old is he? How many brothers has he? How much does he weigh? How much money does his father make?’ Only from these figures do they think they have learned anything about him. (p. 11)

However, teachers expressed that sometimes they cannot emphasize the “essential matters” due to testing requirements and washback from these tests. Both the focus teachers and the surveyed teachers conveyed that they find more value in their classroom assessments than in large-scale assessments. However, surveyed teachers communicated that school administration and district personnel both view large-scale assessments as relatively more valuable tools. Surveyed teachers most frequently cited school, district, and state assessment requirements as the most significant factors that prevent them from making desired changes in the area of assessment for emerging bilingual students. Teachers also cited a lack of power or control as the second most common factor that bars them from making changes in this area. Teachers also communicated internal factors, such as limited time, to make desired changes.

There were insufficient data to make a claim regarding the effect of CLD educational experience as a factor in teachers’ critical use of assessment for emerging bilingual students. While it appears that CLD educational experience has a mostly positive effect on critical use of assessment for the focus teachers, the evidence was less straightforward based on the results from the surveyed teachers. It is understandable that these results were inconclusive, since this was not the main purpose of the study. It would be interesting to see future studies that could survey additional teachers and include, as part of the survey, open-ended items that ask for more specific information about why teachers choose to create or modify their own assessments or why they choose to collaborate with other colleagues or not.
Profile of a Critical User of Assessment

After analyzing the characteristics of teachers who engaged in critical thinking and actions related to assessment for emerging bilingual students, I was able to develop an evidence-based profile of a teacher who is a critical user of assessment for emerging bilingual students.

The critical user of assessment habitually engages in an array of activities related to critical thinking and critical action. Teachers who think critically use assessment information of their emerging bilingual students from multiple measures to adjust their future language instruction to their students’ needs. Critically minded teachers use assessment information to tailor their future content area instruction to their students’ needs as well. Critical users of assessment develop metacognition in their assessment practice to consider the role that language plays in assessment for multilingual learners. Critically minded teachers are careful to avoid using the representativeness heuristic when they formulate their expectations about their emerging bilingual students.

Teachers who are critical users of assessment for emerging bilingual students become involved in certain activities related to the critical engagement component. First, critically minded teachers are intrinsically motivated to engage in critical activities instead of being required to do so. Critical users of assessment focus on instilling critical thinking habits in their own emerging bilingual learners, so that their students might become empowered. Helping students to develop metacognitive skills regarding their progress also serves to foster the development of independent, empowered learners. This approach could have a multiplier effect on teachers’ efforts, since the students they reach could, in turn, positively influence many other people. Critical users of assessment also create or modify their assessments with issues of
language and accessibility in mind. Finally, critically minded teachers engage in formative assessment as a process that comprises part of their regular instructional practice.

Critical users of assessment for emerging bilingual students engage in certain activities if their behavior is not restricted by external factors. Critically minded teachers, for example, have a focus that extends beyond instructional practices driven by washback from large scale assessment. They have a flexible curriculum to meet students’ needs. Critical users of assessment therefore have ample time to regularly engage in activities to foster critical and creative thinking development for their emerging bilingual students. Critically minded teachers are also able to implement culturally relevant, social justice-oriented instruction within the curriculum. They embed language learning opportunities within their instructional practice and take care to promote students’ socio-emotional development. They pay special attention to learning processes instead of focusing almost solely on learning outcomes. Critically minded teachers use their time wisely to be able to plan to habitually incorporate all of these kinds of activities into their practice. Critical users of assessment develop critical thinking skills in their students without a fear of their students becoming critical of them.

*Reflections on the Profile of a Critical Assessment User*

Although the aforementioned traits that characterize the profile of a critical user of assessment are not always explicitly related to assessment, they are still closely tied to the concept. Making a social justice-oriented curriculum a priority, engaging in culturally-responsive instruction, building skills such as critical and creative thinking, and paying attention to socio-emotional development are all activities that are assigned a status of low priority when high-stakes testing practices take precedence over all other types of assessment practice.
Some aspects of the findings do not fit in the profile of a critical user of assessment, yet they are related to the profile. Certain external factors, such as district or state requirements, may act as a barrier to a teacher who aspires to act as a critical user of assessment, yet does not have the freedom to do so. It is necessary for teachers to feel empowered and unrestricted by requirements that might prevent teachers from engaging in critical activities.

Critical users of assessment also value an array of assessment methods that allow for the cultivation and assessment of aspects that escape measurement by standardized tests. This is not to suggest that large-scale tests need to be developed to measure these additional aspects that are currently overlooked. Rather, since these aspects of education are rather ephemeral in nature, I propose that teachers be given the freedom to use more classroom-based, informal assessment practices in conjunction with large-scale testing practices such that they can focus more of their instructional and assessment efforts on these aspects. The consequences of large-scale test results would need to be de-escalated in order to provide opportunities for this shift to occur.

If one cannot measure certain human skills very precisely, it does not mean that they are less important than other skills. It may be elusive to measure any latent trait, from language to cognition to critical thinking skills. Rich context is necessary to help educators grasp the extent to which students may be skilled in any latent trait. Necessary context can be achieved when teachers gather multiple measures of a student’s progress in developing any certain skill, reflect on what they perceive the student to understand, and then push the student forward in their development by providing activities within their zone of proximal development.
Comparing Profiles of Critical Users of Assessment

The theoretical profile of a critical user of assessment was developed earlier in this work in the “Reflection on Critical Thinking and Critical Pedagogy” section of Section Two of the Literature Review to hypothesize the traits that a critical user of assessment for emerging bilingual students might possess. After collecting data through interviews, observations, and a survey, it became apparent that some of the evidence-based characteristics of a critical user of assessment were different than the originally-identified, theoretically-based traits. However, there was some overlap between the theoretically-based and the evidence-based profiles of a critical user of assessment.

Both types of profiles posited that critical users of assessment would select, develop, or modify assessment methods with language in mind to prevent invalid assessment use. Both profiles described how critical assessment users would use assessment methods to promote student learning, although only the evidence-based profile explicitly depicted a broad curriculum encompassing skills such as creative thinking and concepts like culturally responsive instruction. The theoretical profile was general in its reference to student learning. Both profiles suggested that a critical user of assessment would employ assessment methods to inform one’s instructional practice based on the linguistic and content-area needs of one’s students. In this way, formative assessment emerges as an important trait in both profiles. Both profiles described how critical users of assessment would be more aware of the role that language plays in assessment for multilingual learners. They both stressed the development of students’ metacognitive skills, tracking their progress by using assessment tools. Both profiles encouraged the use of multiple measures to develop expectations about students, so that teachers do not make unjustified decisions about their students. Finally, both profiles emphasized the importance of fostering
critical thinking skills in students so that they could access the freedom that accompanies a critical perspective. These characteristics are supported by both theory and evidence in this study, since they were present in both the theoretical and the evidence-based profiles. Therefore, these traits should be more thoroughly examined to determine if they continue to emerge as important characteristics of critical users of assessment in future research.

Implications for Policy

Implication 1: Positioning of teachers as capable of critical thinking. The results of this analysis pose a call to action on the part of policy makers, as one of the major findings of this study is that teachers feel that much of the assessment process is out of teachers’ control. Although the assessment process is an integral part of a teacher’s practice, teachers from this study communicated that there are a host of external forces that prevent them from shaping assessment and instruction in ways they believe to be well-founded for their emerging bilingual students. Almost a decade ago, K. E. Johnson (2006) already identified that washback from large scale assessment had a detrimental impact on teachers’ empowerment (Richard-Amato, 2010). K. E. Johnson claimed that:

[I]t is not surprising that L2 teachers struggle to reject a teach-for-the-test mentality [and] are frustrated by being positioned as managers of curricula than as facilitators of the L2 learning process. (p. 248)

This exemplifies the fact that teachers have become increasingly positioned as managers who are responsible for raising students’ test scores instead of as professionals responsible for developing thoughtful, critical, and bilingual members of society. As this trend continues, teachers are going to continue to feel less empowered in their roles. Teachers’ roles have evolved greatly since the NCLB legislation was passed, and it has become more difficult for teachers to play both the
“numbers game” role as well as the broader role of raising thoughtful future members of society.
If students continue to be considered in terms of numbers by policy makers, teachers have less ability to humanize and empower their students. All students are labeled as unsatisfactory when they do not make adequate gains on their test scores from year to year. In Colorado, students who fail to achieve a certain expected reading level are deemed as having a “significant reading deficiency.” If a student is an emerging bilingual, the labels do not stop there. Students are labeled “Non English Proficient” or “Limited English Proficient” based on their test scores.

As participants have testified in this study, educators sometimes struggle to showcase their students’ progress and convince them that they are making progress despite the negative labels assigned to them. However, teachers of emerging bilinguals are rated by the same metrics as their students, so they also face negative repercussions when their students do poorly on tests that were not designed for them (Gottlieb, 2000). They stand to lose their position or be humiliated by the public display of test scores (Los Angeles Times, 2015). They do not feel trusted because of the heightened emphasis on accountability since the turn of the century. Teachers are therefore not in a position to empower their students when their own livelihood is at stake. In order for teachers to facilitate students in breaking free from the negative labels and stereotypes that accompany these language- and performance-related labels, teachers need to be critical users of assessment. However, as demonstrated by the survey results, many teachers feel restricted by legislative requirements and policies surrounding assessment.

**Implication 2: Increased importance of critically examined, linguistic- and content-based classroom assessment methods.** Assessment-related policy should value teachers as critical users of assessment by placing importance on teachers’ classroom assessment methods. More importantly, assessment-related policy should recognize and value teachers’ perspectives
regarding the performance of students with whom they work each day. If schools aspire to create thoughtful, critical bilingual society members then it is imperative to recognize their development in a broader sense: ethically, culturally, and socio-emotionally as well as academically. This recognition is impossible as long as large-scale test scores serve as the leading, and sometimes sole, metric to determine student and teacher progress over the course of a year. Teachers need to be given the space to ensure their assessment methods reflect all of the ways in which we would like students to develop. In doing so, policy would need to allow for multiple measures of student progress that recognize informal assessment methods. This entails a reduced emphasis on the importance of large-scale test scores and an increased focus on classroom-based assessments.

Legislation would need to allow for more control of the assessment process to be placed in the hands of teachers and students in the classroom. These changes would serve as a shift from assessment serving a mainly punitive role to serving a more informative, empowering role for teachers and students to use as they see fit. As assessment currently is practiced, teachers do not have the opportunity to “use” large-scale assessment hardly at all for linguistic- or content-based purposes. For test security purposes, teachers do not have access to the content of large-scale test items prior to testing. They are instructed not to look at students’ exams during test administration. Individual student responses to test items are not shared with teachers to maintain the secrecy of the reading passages and the test items. They act as proctors, and in preparation for the tests they are merely managers of information to be transmitted to students. The assessment process could be employed to humanize students who have been labeled with deficiencies and relegated to a low status with corresponding levels of self-esteem that serve to perpetuate social inequities in the world outside the school doors. In order to strive for social
justice for all learners, particularly emerging bilinguals, it is imperative to assign importance to classroom-based assessment methods that incorporate informal measures and recognize the whole student’s development.

*Implication 3: Professional development for teachers to learn to act as better critical users of assessment.* With fewer large-scale testing sessions necessary, excess funding could be allocated to provide professional development programs to support teachers in developing their own formative and summative assessments. This professional development would include repeated coaching sessions so as to provide training specific to the needs of each teacher. As multiple focus teachers in the study stated, it is widely recognized that teachers need to differentiate for students. It is equally important for professional development to attend to the varying needs of teachers as well (Alma, personal communication, April 28, 2014; Valera, personal communication, April 9, 2014).

*Implication 4: Time allotted for teachers to act as critical users of assessment.* Teachers also need the time to be able to develop their own assessment methods to complement large-scale assessments. These classroom-based assessment methods have the potential to be specific to the linguistic and developmental needs of students. They can be culturally relevant, authentic to the students’ lived experiences outside of school, and they can emphasize the process of learning. These assessment methods can serve as tools to aide teachers in making education a transformative experience for students. These assessment methods also have the potential to serve as part of a larger overall approach to steer away from school’s traditional role in social reproduction of inequities in large society.

*Implication 5: Need for policy makers to act as critical users of assessment for emerging bilinguals.* Language assessment issues in the political realm are closely related to System I and
II concepts. In Kahneman’s (2010) chapter on “the Illusion of Validity,” he concludes that “errors of prediction are inevitable because the world is unpredictable” (p. 220). He discusses how economic and political ‘pundits’ who make a living off of their predictions should not be trusted. They may have a high level of confidence about their predictions, but according to a study published in a book by Tetlock (2005), experts perform worse than if someone randomly predicted political and economic trends in an even distribution across the options available. However, “illusions of validity and skill are supported by a powerful professional culture. We know that people can maintain an unshakable faith in any proposition, however absurd, when they are sustained by a community of like-minded believers” (Kahneman, 2010, p. 217).

Kahneman urges people not to trust indicators or tests that are not shown to be valid. It is important to recognize that “reality emerges from the interactions of many different agents and forces, including blind luck, often producing large and unpredictable outcomes” (p. 220).

It is important to recognize that it is not only financial and political pundits or Wall Street stock traders who place excessive confidence in their predictive skill when elements of luck and randomness are actually responsible for the outcome of a situation. In the education profession, the business concept of ‘Value-Added Modeling’ (VAM) has taken precedence in the evaluation of teachers. This modeling is said to predict a teacher’s performance over the year to yield a valid evaluation, based on their students’ standardized test scores at one point in the year. These evaluations are important because they determine who retains their job year to year as well as who receives financial bonuses. Sometimes teachers are even publicly recognized or shamed for their students’ test score performance in the media.

However, evaluations based on student achievement are not always dependable, particularly for linguistically diverse groups. There are many factors outside of a teacher’s
control that contribute to their value-added modeling score. As the American Statistical Association (2014) writes: “Most VAM studies find that teachers account for about 1 percent to 14 percent of the variability in test scores, and that the majority of opportunities for quality improvement are found in the system-level conditions” (p. 2). The parallels to the financial world are strong, and there are lessons to be learned. It would be prudent for professionals in both arenas to look at the evidence, incorporate more System II thinking, and question the assertions made about teacher performance or financial events.

Half of the focus teachers in this study were able to use System II thinking to arrive at these questions (specifically Valera, Serena, and Alma), yet they do not have an arena to voice these concerns. Teachers in the study wanted to know how they could incite change in this regard. Unfortunately, most of these issues lie outside of teachers’ nexus of control. The power to change lies in the hands of policy makers at the district, state, and federal levels. Sustained grassroots efforts have potential to change the minds of policy makers, but assessment determinations are ultimately made by legislators and administrators who interpret assessment-related legislative policies.

Limitations

I have identified five limitations in this research. The first limitation is mainly related to the small scale of this study. Since this study is a dissertation and was not allocated the resources that might be afforded to a larger grant-funded project, the sample size for the study was meager. If I had access to additional time and resources, I would have interviewed, observed, and surveyed more participants. I was very fortunate to have been awarded a dissertation grant to compensate the teachers who participated in this study. However, I still only managed to gather 72 survey participants, 12 of whom did not meet the required demographic criteria. Ideally, I also
would have incorporated all of the data collected for this dissertation. I already had access to a large amount of data by using information collected from only six of the ten teachers I interviewed and observed. I made use of only the first interview out of two interviews with each of the six focus teachers since the first interview was so full of rich stories and ideas that it provided me with more than enough data to analyze. However, in the future I would like to incorporate both interviews from all teachers in the study.

Second, the participants for the survey, the interviews, and observations were not randomly selected. The small scale of this study limited the potential to conduct a randomized study on teachers in Colorado, who have many competing interests on their time. Given that the goal of this study was to explore in depth how teachers of emerging bilingual students might act as critical users of assessment, I thought it important to find teachers who would be willing to invest hours of their time to participate in interviews and observations with me rather than to randomly select a larger sample of teachers with whom I could not spend as much time. I was still able to elicit survey responses from 30 participants in each survey group (CLD and non-CLD) who satisfied the demographic criteria I set in order to provide a larger sample of teachers’ perspectives. Also, the results of this study are not intended to be generalizable to the overall population of teachers of emerging bilingual students in Colorado. Rather, this study is meant to highlight potential themes that can be investigated further for the purposes of generalizability.

Third, I was not able to develop interview and survey items that could control for the natural human inclination to slightly overstate one’s abilities in a skill area. Respondents would like to appear as capable as possible in the context of an interview, and people tend to consider themselves as being highly competent professionals in their practice (Atir et al., 2015; Lakin & Wallace, 2015; Lee, Hart, Cuevas, & Enders, 2004; Lee, Penfield, & Maerten-Rivera, 2009;
Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; Randall & Fernandes, 1991). Though some of the teacher participants asserted more confidence in their abilities than others and were therefore more likely to overestimate their skills, it was not possible within the context of the study to determine if this overestimating phenomenon was truly at work. Therefore, no actions were taken to adjust for this.

Fourth, the coding procedure was rather specific to the context of the conceptual framework and therefore I was the sole coder of the data analyzed for this project. Ideally, I would establish inter-rater reliability with a team of researchers with whom I would code the data. However, my coding procedures were consistent and I discuss them clearly in the methods section.

Finally, it is quite challenging to understand and identify aspects of assessment that teachers consider when they assess emerging bilinguals. This study is particularly complicated in realistic school settings, where many factors influence the way in which teachers think and act. Thus, the study of teachers’ critical use of assessment is a difficult endeavor. It is necessary to acknowledge that there are always limitations to measurement of the human mind and its motivations. Teachers’ thought processes are latent variables, which cannot be measured in a direct manner. As Mlodinov (2008) remarked in his book about randomness in measurement, “It is one of those contradictions in life that although measurement always carries uncertainty, the uncertainty in measurement is rarely discussed when measurements are quoted” (p. 129). This is a fundamental understanding that I attempt to acknowledge as clearly as possible in my own work. It is a privilege to have the opportunity to study human thought and action. I respect the fact that there will be error in my own evaluation of teachers’ critical use of assessment, and I have been as clear as possible about the presence of error in my own work.
**Future Research**

This study sought to explore critical concepts in the intersection of the second language acquisition and assessment fields. Specifically, I attempted to identify the extent to which teachers are critical users of assessment for emerging bilingual students, and to determine the factors that prevent teachers from acting as critical users of assessment. Since this study is intended to serve as an initial exploration into this inquiry, this area is ripe with opportunities to conduct further research in more depth. It would be beneficial to students for researchers to develop a more nuanced understanding of how System I, System II, and CE component activities are related to each other. It would be helpful to investigate if these components share a sequential relationship. Specifically, it would be interesting to know when and why CE component activities become activated. This would shed light on how to help teachers to engage more habitually in CE component activities.

Future research could focus on conducting a more thorough analysis of teachers and their critical use of assessment for emerging bilingual students by gathering a randomized sample that is representative of teachers of emerging bilinguals in Colorado or even nationwide. A team of coders could further refine my coding framework by carefully defining the activities in each component such that coding reliability could be established. This could help to solidify the claims made in the results chapter and to make them more generalizable to a larger population of teachers of emerging bilingual students.

If the claims are indeed generalizable to the population of teachers following that study, a research project focused on providing professional development for teachers, by teachers, should be created. Researchers would identify teachers who closely fit the profile of a critical user of assessment to take part in planning the professional development series. This series would help
teachers to develop their classroom assessment methods from a critical perspective, with an attentive eye on issues of language. This training would emphasize the importance of teachers’ role as educators to develop critical thinking skills in their students. It would also stress that it is imperative for teachers to engage in critical activity when they recognize problems in their field. Collective action is necessary for teachers to create positive change in assessment for emerging bilinguals, since much of the power lies outside of individual teachers’ nexus of control. Teachers would be an integral part of the professional development planning process and then implement the sessions with groups of teachers.

Conclusion

When I began to write this dissertation, I wrote with the intention of inspiring educators to become more critical and vocal about the issues they encounter regarding assessment of emerging bilingual students. I also wrote to shed light on the factors that act as barriers to teachers in their ability to voice their concerns about these issues. I had a certain image in my mind of what critical action might potentially look like as a response to the “numbers game” approach to education. I imagined that, when and if teachers did not agree with a certain requirement, teachers might take political action and organize a grassroots resistance. This kind of political engagement does occasionally occur, although I was unable to capture it in my study.

Now that I have spent many hours reflecting on the results of this study; however, I realize that teachers do not necessarily need to take action outside of the classroom in order to act as critical users of assessment for emerging bilingual students. Clearly, it is optimal for teachers to act as critical users of assessment in a variety of ways, from organizing political action to analyzing the linguistic aspects of assessment methods with their colleagues. Yet
teachers engage in many activities even within their own classrooms that fall under the description of critical assessment use.

Initially, I underestimated the importance of one of these activities. Teachers’ actions to inspire critical engagement in their students can be very powerful. Teachers can develop critical consciousness in their students by teaching them to metacognitively monitor and assess their own progress with regard to language and content. By doing this, teachers act as critical users of assessment by helping their students become independent learners who are capable of breaking free from the negative labels they encounter in society. They can involve their students in analyzing and modifying language assessment methods based on students’ expressed needs, thereby engaging their students to be thoughtful and critical language learners. Teachers can work together with their emerging bilingual students to strive to develop habits of critical thinking and engagement in the classroom and beyond. It is in this collaboration that teachers become everyday heroes, doing the work of empowering emerging bilingual students so that they can actualize their hopes and dreams for the future.
References


Nguyen-Le, K. (2010). *Personal and formal backgrounds as factors which influence linguistic and cultural competency in the teaching of mathematics, UMI 3419499*. University of Colorado at Boulder.


The Colorado Department of Education. (2013). Adding an endorsement to a valid teacher or special services provider license -. Retrieved from http://www.cde.state.co.us/cdeprof/licensure_addendorsement_info


Appendix A.
Survey for Teachers

Default Question Block

Welcome to the Teacher Use of Assessment Survey.

Thank you for providing your input on this survey. This survey is the first part of a study that inquires about teachers' use of assessment with emerging bilingual students. Your responses will help to build an understanding of how teachers use assessment in their classrooms and how to better support teachers in this endeavor. This survey should take about 15 minutes to complete.

You have been selected to participate in this survey because you have been identified as a current teacher of English Language Learners (ELLs) in Colorado. I hope that I can count on your participation in this survey.

This survey is confidential. Serious efforts will be made to limit the use and disclosure of your personal information, including research study records, to people who have a need to review this information. We cannot promise complete secrecy. Organizations that may inspect and copy your information include the IRB and other representatives of this organization. Your completed survey form and this consent form will be retained in a locked cabinet for security purposes. Only the research team will have access to this information. The surveys and consent forms will not be retained for longer than three years.

Participation in the survey is completely voluntary. We urge you to participate in the survey—your responses and experiences will be invaluable in helping educational professionals better understand the way in which teachers use assessment with emerging bilingual students so that we can find appropriate and helpful ways to support teachers in their practice.

Should you decide to participate in this survey, you will be entered into a drawing to win $100. If you are a winner of the drawing, you will be contacted by email at the address you provide at the end of the survey.

To acknowledge your consent to participate in this survey, select the option below and click "NEXT."

- I have read and understand this consent form. I choose to participate in this survey.
- I do not choose to participate in the survey at this time.

Are you currently teaching in the state of Colorado in a P-12 public, private, or charter school setting?

- Yes, in a public school
- Yes, in a private school
- Yes, in a charter school
- No, (this includes if you are currently teaching outside of Colorado, serving as a short-term substitute in different classrooms, or working less than a 0.5 full time position)

On average, what percentage of students in your class(es) have been English Language Learners (ELLs) over the past 3 years?

- 0%
- 1 - 25%
- 26 - 50%
- 51 - 75%
- 76 - 100%

How many academic years of full time teaching experience have you completed?

- less than 1 year
1 year
2-3 years
4-5 years
6-9 years
10-14 years
15-19 years
20+ years

Are you trying to obtain a Culturally and Linguistically Diverse endorsement to teach ESL or Bilingual Education or do you already hold this endorsement?
- Yes, I am currently working towards my CLD endorsement in the state of Colorado.
- Yes, I currently hold a CLD endorsement in the state of Colorado.
- No, I do not have (nor am I working towards) a CLD endorsement.

Which institution are you attending in order to obtain the CLD endorsement in Colorado?

Which institution did you attend in order to obtain the CLD endorsement in Colorado?

In which school district (or county if you do not teach in a public school) of Colorado do you currently teach?

What is your sex?
- Female
- Male

What is your current age?
- 18-25
- 26-35
- 36-45
- 46-55
- 56+

Which grade(s) do you currently teach? Please select all that apply.
- Pre-school
- Kindergarten
- Grade 1
- Grade 6
- Grade 7
- Grade 8
TEACHERS AS CRITICAL ASSESSMENT USERS FOR EMERGING BILINGUALS

Grade 2
Grade 3
Grade 4
Grade 5

Grade 9
Grade 10
Grade 11
Grade 12

What subject area(s) do you currently teach? Please select all that apply.

- Early Childhood Generalist
- Elementary Generalist
- Language Arts
- Math
- Science
- Social Studies
- Reading
- World Language(s)
- Music
- Art
- Physical Education
- Library
- English as a Second Language
- Special Education
- Speech
- Bilingual Education Program
- Other - please describe: ________________________________

Which type of bilingual program does your school implement?

- Early Exit Transitional Bilingual Education
- Late Exit Developmental Bilingual Education
- Dual Language or Two-Way Bilingual Program
- Heritage Language Program
- Other, please explain: ________________________________

Which type of English as a Second Language program does your school implement?

- Pull-Out Model
- Push-In Model
- ESL Period
- ESL Resource Room
- Structured English Immersion
- Other, please explain: ________________________________

How long have you been teaching this subject/these subjects?

- 1 year or less
- 2-3 years
- 4-5 years
- 6-9 years
10-14 years
○ 15-19 years
○ 20+ years

What subject area(s) are you certified to teach in Colorado? Please include areas in which you have an endorsement. Please select all that apply.

- Early Childhood Generalist
- Elementary Generalist
- Language Arts
- Math
- Science
- Social Studies
- English as a Second Language (CLD endorsement)
- Reading (endorsement)
- World Language(s)

- Music
- Art
- Physical Education
- Library (endorsement)
- Special Education
- CLD Bilingual (endorsement)
- Speech
- None
- Other - please describe: 

Do you speak any languages other than English?
○ Yes
○ No

Which languages do you speak, and at which levels of proficiency? Please begin by listing your first (native) language and continue with the other languages that you speak.

Note: If you learned two languages at the same time, please indicate that in parentheses after the language name. For example, if I learned Spanish and English both at home in my early childhood, I would list "Spanish (concurrent)" and "English (concurrent)" in the spaces for my first and second language.

<table>
<thead>
<tr>
<th>Language</th>
<th>Beginner</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Proficient</th>
<th>Native Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>First (Native)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fourth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The purpose of this survey is to develop an understanding of actual teacher practices, so please be honest.

I receive the most helpful information about my ELLs' progress towards CONTENT AREA goals from:
○ Homework
○ Classroom tests
Quizzes

- In-class activities
- Students’ questions / conversation in class
- Large-scale tests

I receive the most helpful information about my ELLs’ progress towards LANGUAGE PROFICIENCY goals from:

- Homework
- Tests
- Quizzes
- In-class activities
- Students’ questions / conversation in class
- Large-scale tests

When I receive helpful information about my ELLs’ progress in a class, I MOSTLY:

- Use it to help me in my grading and reporting of student progress
- Make inferences from it to help me know what to do for students in future language instruction
- Make inferences from it to help me know what to do for my students in future content area instruction
- Use it to get a good idea of where my students are in their learning
- Use it in some other way:

I MOSTLY think assessment is:

- a practice that takes time away from my instruction
- a practice that is a part of my instruction on a unit level
- a practice that is a part of my daily instruction
- a way to learn more about my students and what they can do, but it’s hard to know how to use that information
- something else:

For the purposes of the next part of this survey, please consider assessment as “a comprehensive process of planning, collecting, analyzing, reporting, and using information with students over time. Assessment can include tests as well as projects, anecdotal information, and student self-reflection” (Gottlieb, p. 86).

In general, I create my own assessments.

<table>
<thead>
<tr>
<th>Always (6)</th>
<th>Frequently (5)</th>
<th>Usually (4)</th>
<th>Sometimes (3)</th>
<th>Rarely (2)</th>
<th>Never (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I change or modify parts of assessments that I do not create myself.

<table>
<thead>
<tr>
<th>Always</th>
<th>Frequently</th>
<th>Usually</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
How do you change or modify the assessments?

In my interpretation, I think my classroom assessments are important to:

<table>
<thead>
<tr>
<th></th>
<th>Always</th>
<th>Frequently</th>
<th>Usually</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>Me, as a teacher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My students’ parents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My principal / admin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My district</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In my interpretation, I think large-scale assessments (at the district, state, or national level) are important to:

<table>
<thead>
<tr>
<th></th>
<th>Always</th>
<th>Frequently</th>
<th>Usually</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>Me, as a teacher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My students’ parents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My principal / admin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My district</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

My colleagues and I collaborate to plan our assessment methods.

<table>
<thead>
<tr>
<th></th>
<th>Always</th>
<th>Frequently</th>
<th>Usually</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
</table>

Please briefly describe why you selected that answer.

When it comes to designing and implementing high-quality assessment methods to measure how my ELLs are progressing with a unit of study, I feel:

- very prepared and confident
- mostly prepared and confident
- somewhat prepared and confident
- I could really use the help of a class or professional development for this
If you could change anything about assessment used with ELLs (either large-scale assessments or classroom-based assessments), would you?

- Yes
- No

What would you change, and why?

Is there anything that prevents you from making changes? If so, what?

Please provide your contact information, only so that I may contact you if you are selected as a winner of the raffle.

Name: 

Email address: 

Block 1

I'm sorry. From the information you've provided, you do not appear to fit the characteristics of the target population for this survey. Your survey will now end.
Appendix B.
Interview protocol 1: General Interview

Qualtrics-administered demographic component:

Are you currently teaching in the state of Colorado in a P-12 public, private, or charter school setting?

- Yes, in a public school
- Yes, in a private school
- Yes, in a charter school
- No, (this includes if you are currently teaching outside of Colorado, serving as a short-term substitute in different classrooms, or working less than a 0.5 full time position)

On average, what percentage of students in your class(es) have been English Language Learners (ELLs) over the past 3 years?

- 0%
- 1 - 25%
- 26 - 50%
- 51 - 75%
- 76 - 100%

How many academic years of full time teaching experience have you completed?

- less than 1 year
- 1 year
- 2-3 years
- 4-5 years
- 6-9 years
- 10-14 years
- 15-19 years
- 20 + years

Are you trying to obtain a Culturally and Linguistically Diverse endorsement to teach ESL or Bilingual Education or do you already hold this endorsement?

- Yes, I am currently working towards my CLD endorsement in the state of Colorado.
- Yes, I currently hold a CLD endorsement in the state of Colorado.
- No, I do not have (nor am I working towards) a CLD endorsement.

Which institution are you attending in order to obtain the CLD endorsement in Colorado?
Which institution did you attend in order to obtain the CLD endorsement in Colorado?

In which school district (or county if you do not teach in a public school) of Colorado do you currently teach?

What is your gender?
- Female
- Male

What is your current age range?
- 18 - 25
- 26 - 35
- 36 - 45
- 46 - 55
- 55 +

Which grade(s) do you currently teach? Please select all that apply.
- Pre-school
- Kindergarten
- Grade 1
- Grade 2
- Grade 3
- Grade 4
- Grade 5
- Grade 6
- Grade 7
- Grade 8
- Grade 9
- Grade 10
- Grade 11
- Grade 12

What subject area(s) do you currently teach? Please select all that apply.
- Early Childhood Generalist
- Elementary Generalist
- Language Arts
- Math
- Science
- Social Studies
- Reading
- World Language(s)
- Music
- Art
- Physical Education
- Library
- English as a Second Language
- Special Education
- Speech
- Other - please describe: [ ]
Which type of English as a Second Language program does your school implement?
- Pull-Out Model
- Push-In Model
- ESL Period
- ESL Resource Room
- Structured English Immersion
- Other, please explain: 

Do you teach in a school with a bilingual program?
- Yes
- No

Which type of bilingual program does your school implement?
- Early Exit Transitional Bilingual Education
- Late Exit Developmental Bilingual Education
- Dual Language or Two-Way Bilingual Program
- Heritage Language Program
- Other, please explain: 

How long have you been teaching in your current subject area(s)?
- 1 year or less
- 2-3 years
- 4-5 years
- 6-9 years
- 10-14 years
- 15-19 years
- 20+ years

What subject area(s) are you certified to teach in Colorado? Please include areas in which you have an endorsement. Please select all that apply.
- Early Childhood Generalist
- Elementary Generalist
- Language Arts
- Math
- Science
- Social Studies
- English as a Second Language (CLD endorsement)
- Music
- Art
- Physical Education
- Library (endorsement)
- Special Education
- CLD Bilingual (endorsement)
- Speech
Are you currently working towards any other teaching certifications or endorsements in Colorado?

- Yes
- No

Which certifications or endorsements are you working towards? Please select all that apply.

- Early Childhood Generalist
- Elementary Generalist
- Language Arts
- Math
- Science
- Social Studies
- English as a Second Language (CLD endorsement)
- Reading (endorsement)
- World Language(s)
- Music
- Art
- Physical Education
- Library (endorsement)
- Special Education
- CLD Bilingual (endorsement)
- Speech
- Other - please describe:

Do you speak any languages other than English?

- Yes
- No

Which languages do you speak, and at which levels of proficiency? Please begin by listing your first (native) language and continue with the other languages you speak.

Note: If you learned 2 languages at the same time, please indicate that in parentheses after the language name. For example, if I learned Spanish and English both at home in my early childhood, I would list “Spanish (concurrent)” and “English (concurrent)” in the spaces for 1st and 2nd language.

<table>
<thead>
<tr>
<th>Language</th>
<th>Beginner</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Proficient</th>
<th>Native Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>First (Native) Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fourth Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other - please describe:
Orally Administered Interview Component:

Context:
The goal of this study is to learn about how you use assessment in your classroom for all your students, especially emerging bilingual students. This interview will ask questions about your teaching practice in regards to assessment, about your teaching and about your environment in general.

Understanding teachers’ conceptions of assessment

1) In any given unit, how do you know if all of your students are learning and making progress towards content area goals? (If you do very different things in different content areas, please briefly describe these differences. You can give specific examples as you see fit.)

2) In any given unit, how do you know if all of your students are learning the language of the relevant content area? (Don’t ask leading questions about academic language.)

   Do you use any other information to better understand your students’ progress?

3) How do you determine students’ grades at the end of a marking period?

   (If this answer is different from their answer to “How do you know all your students are learning?” then ask:

   “Why is your grading strategy different from how you keep track of your students’ learning on a daily basis?”)

4) What do you think of when you think of the word “assessment”? What “counts” as assessment? (If you were going to give a definition, what counts as assessment in your book?)

   Interviewer will define assessment according to Gottlieb and then clarify different types of assessment at this point:

   “To ensure consistency in our language for the purposes of this interview, I’m going to clarify what I mean when I use the term “assessment”. There is no right answer – I just want to clarify this because assessment is a term that can be used in various ways and have different connotations. Assessment can be defined by Margo Gottlieb as “a comprehensive process of planning, collecting, analyzing, reporting, and using information with students over time. Assessment can include tests as well as projects, anecdotal information, and student self-reflection” (Gottlieb, 86). Assessments can be formal (planned) or informal (unplanned) in nature. It can also include instruments or activities. Assessment can also be for a summative, end-of-unit purpose, or it can be for formative purposes as you progress through a unit of study with your class.

5) What does the term “English Language Learner” or “ELL” mean to you? What do you think the term indicates about a student?
“English Language Learner (ELL) is also a term that can have various connotations. To ensure consistency in our language for the purposes of this interview, I’m going to clarify what I mean when I use the term “ELL”. According to Margo Gottlieb, an ELL can be defined as a “linguistically and culturally diverse student qualified for support services.” Students are typically identified for support services through an assessment process.

Do you have any questions before we go on?

*Teachers’ assessment practices*

6) Please tell me about other kinds of assessment methods you use in your class, if you haven’t mentioned or discussed them already.

*If this is not addressed by the teacher in the course of the previous conversation, ask:*

Do you consider end-of-unit formal assessments summative, formative, or a combination of both?

7) What goals or objectives do you have for your students when you assess their progress?

Do you integrate language development and content area learning goals in your classroom assessments?

- *If yes:* How do you integrate language development and content area learning goals into your classroom assessments?

- *If no:* Can you tell me how you determine which goals you have for students in your assessments?

8) When you devote time to formal assessment in your classroom, approximately what percentage of this time are the assessments self-created?

*If reply is above 50%:* How do you create your own assessments? *WHY?*

9) When you devote time to externally developed formal assessments, approximately what percentage of this time do you change or modify them?

*Unless it’s 0%:* How do you modify formal assessments that are not self-created? *WHY?*

10) When you devote time to assessment in your classroom, what percentage of this time are your students active participants in the assessment process?

*If above 50%:* How are your students active participants in the assessment process? 
*If below 50%:* Please tell me more about who participates in your classroom assessment practice, and how do they participate?
11) When you devote time to assessment in your classroom, what percentage of this time do you give everyone the same classroom assessments, including ELLs?

If above 50%: Can you tell me about how you make the decision to give everyone the same assessments or not?
If below 50%: Can you tell me about how you make the decision to differentiate your assessment practice among various students?
If 0%: Why? Or just say “OK”.

12) To what extent do you familiarize yourself with new assessments that your students will be evaluated by?

If “a great deal” or something similar: How do you familiarize yourself with these new assessments?
If “not a lot” or something similar: Why?

13) Do you feel prepared to design high-quality assessments that help you to understand how all of your students are progressing with the language and content in a unit of study?

Why did you rate yourself the way you did?

14) When you devote time to formal assessment, what percentage of this time do you reference a textbook or workbook to copy material directly?

If above 50%: Can you tell me more about what you take from textbooks or workbooks and how you make the decision to use it?
If below 50%: Can you tell me more about how you make that choice?

15) Do you regularly set aside time to reflect on your assessment practices?

If yes: What aspects do you consider (or take action on) when you reflect on these practices?
If no: Can you tell me about what influences your decision?

The school and the community

16) Do you feel there is a characteristic way of engaging in assessment practices in this school? If yes, how does that impact your own assessment practices?

17) Does the administration ask that you conduct some assessment practices in a certain way? If yes, how does that impact your classroom assessment practices?

18) Do you know what assessment methods other teachers in your school use?
How do you know? (if they mention collaboration, move to next question)
Do you observe other classrooms?
19) When you plan your assessment methods, what percentage of this time do you and your colleagues collaborate?

   Can you tell me about what influences your decision to collaborate or not with your colleagues about assessment methods?

   Do you collaborate with your colleagues on topics other than assessment?

20) Do you communicate with other teachers or specialists on a regular basis?

   If yes – what do you communicate about?

   Do you communicate about assessment?

21) Do you involve parents in the assessment process in any way?

   If yes, How do you involve them?

The larger assessment environment

22) Do large-scale tests, such as TCAP, influence your use of classroom assessment in any way?

23) Do you think that large-scale tests reflect your students’ progress in your class?

   If yes, can you describe how you came to that conclusion?

   If no, what is the discrepancy, and how did you come to that conclusion?

24) At what point do you think emerging bilingual students are no longer impacted by their language proficiency when taking assessments?

   Can you describe to me how you came to that conclusion?

25) If you had the power to change anything about the assessments used with ELLs (either large-scale assessments or assessments in your school), would you change anything? If so, what would you change?

26) If you answered yes to the previous question, what holds you back from being able to make change?
Appendix C.
Interview Protocol 2: Assessment Artifact Interview

Context:
The goal of this study is to learn about how you use assessment in your classroom for all your students.

Introduction:
We are trying to identify all of the actions that you take and aspects of assessment that you consider throughout the process of using assessment tools in your classroom. This interview will focus on the particular classroom assessment that you brought to share today. Therefore, please be as thorough and specific as possible when you are discussing how and why you used the current assessment.

Questions:

(1) Please describe the purpose of this assessment.
   - Will the assessment be used by your students? How?\(^1\)
   - Will the assessment be used to serve you and your instruction? How?
   - Will the assessment be used by students’ parents? How?
   - Will the assessment be used by your administration? How?

(2) Did you select this assessment from a published source, like a textbook, did you create the assessment on your own or with colleagues, or did you combine some of your ideas with some published ideas?
   
   *If they selected the assessment from a published source:*
   - Did you change or modify anything about the assessment before you gave the assessment to your students?

Now I’m going to ask you about the steps you took when you used this assessment in your class, from the beginning to the end.

(3) Please describe in detail the steps you took when you administered the assessment to your students.
   
   *If not mentioned in their description:*
   – Did you administer the assessment to all students in the same way? If yes, stop.
   - If not, what accommodations did you provide?
   - Why or why not?
   - If you provided accommodations, do you think they were effective?
   - Why or why not?

\(^1\) An indent and a dash indicate a probing question, which will not be asked if the respondent answers it in the process of answering the main question posed to them.
(4) Please describe in detail the steps you took when you scored the assessment.

If not mentioned in their description:
- How did you determine the scoring method?
- If applicable, how did you decide if you were going to curve the scores or not?
- Was it easy or difficult to remain consistent in your scoring?
- If not, why not?

(5) Please describe in detail the way in which you interpreted this assessment and any methods you used to help you to interpret the assessment.

- What led you to interpret the assessment in this way?
- Did you notice any trends in the assessment results?
  - If so, what were the trends?
  - Why do you think these trends appeared?
- Do you think the assessment was accessible to all your students in terms of its content?
  - How?
- Do you think the assessment was accessible to all your students in terms of its format?
  - How?

(6) Please find assessments that three of your ELLs completed (or all of your ELLs, if there are fewer than three in the class).

For each student’s assessment:
- Please describe how you interpreted their assessment result.
- How did the student’s assessment result compare to your expectation of their performance on the assessed content?
- (if there is a discrepancy between the assessment results and the expectation of the student’s understanding) Why do you think there is this discrepancy between the assessment results and your expectation of their performance?
- Will this impact your further instruction in any way? If yes, how?

(7) Please find assessments that three of your non-ELLs completed.

For each student’s assessment:
- Please describe how you interpreted their assessment result.
- How did the student’s assessment result compare to your expectation of their performance on the assessed content?
- (if there is a discrepancy between the assessment results and the expectation of the student’s understanding) Why do you think there is this discrepancy between the assessment results and your expectation of their performance?
- Will this impact your further instruction in any way? If yes, how?

(8) Did you make any decision(s) about your students’ progress in understanding your content or language objectives based on this assessment? If so, what decision(s) did you make?
- Will the results of this assessment influence your teaching practice?
  - If so, how?
- Will the results of this assessment influence your understanding of your students’ progress?
  - If so, how?
- Will the results of this assessment influence your understanding of your ELLs’ progress specifically?
- If so, how?

(9) Do you think the assessment was reasonably valid (i.e. accurate in telling you what you wanted to know about each students’ understanding of the evaluated content)? In other words, do you think the assessment was able to tell you what the students understood about the content in the same way?
- Can you tell me how you came to that conclusion?
- Do you think the language used on the test influenced the accuracy of the assessment?
- Can you tell me how you came to that conclusion?

(10) Do you think this assessment successfully served its purpose?
- Can you tell me how you came to that conclusion?
- Did the assessment serve your students? How?
- Did the assessment serve you and your instruction? How?
- Did the assessment serve students’ parents? How?
- Did the assessment serve your administration? How?

11) Is there anything you would change about this assessment, now that you have seen the results? If so, what would you change?
Appendix D.
Code book procedure based on research questions

RQ 1. To what extent do teachers of emerging bilingual students think critically about assessment?

**System II**

<table>
<thead>
<tr>
<th>Concept / Word</th>
<th>General Definition (all in verb form)</th>
<th>Contextualized Definition (Please note: These definitions are not exhaustive. Do not consider these definitions as limiting to other possibilities that may not be mentioned.)</th>
<th>Example in Context with Respect to Assessment with Emerging Bilingual Students</th>
<th>Example from data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compare</td>
<td>To relate or liken two concepts to each other according to the aspects they have in common</td>
<td>When a teacher detects correspondences between two ideas, objects, and the like, in their curriculum or in their students’ progress.</td>
<td>Compare the Common Core standards and the WIDA standards to see how expectations of language and content fit together.</td>
<td>Common Core is like a new fad, it’s like everyone just got twitter. OK remember how crazy that was? That’s what this is like. Every state is adopting the Common Core, and now every state is on the same level, we’re all measuring on the same thing, (Adele, Q7)</td>
</tr>
<tr>
<td>Contrast</td>
<td>To find the differences in two concepts based on the aspects they have that are dissimilar</td>
<td>When a teacher detects differences between two ideas, objects, and the like, in their curriculum or in their students’ progress.</td>
<td>Contrast the Common Core standards and the WIDA standards to see how expectations of language and content do not overlap.</td>
<td>At least the TCAP/CSAP era, it would show me areas of weakness. (Dina, Q4)</td>
</tr>
<tr>
<td>Action</td>
<td>Description</td>
<td>Teacher Activity</td>
<td>Example</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>-----------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Analyze</td>
<td>To examine logic through reading or listening and scrutinizing the points that have been communicated</td>
<td>When a teacher breaks material into its constituent parts and determines how the parts relate to one another and to an overall structure or purpose, such as for the purposes of assessing a student and identifying if their misconceptions should be attributed to language or to content.</td>
<td>To examine student work to see how their progress in language and in content understandings is developing over time.</td>
<td></td>
</tr>
<tr>
<td>Interpret</td>
<td>To figure out or explain the meaning of, to present in understandable terms</td>
<td>When a teacher changes from one form of representation to another, for example, in order to form an accurate understanding of student progress.</td>
<td>Interpret student assessment results to infer students’ understanding about particular topics and the language surrounding those topics.</td>
<td></td>
</tr>
<tr>
<td>Evaluate</td>
<td>To determine the condition of through careful study or appraisal</td>
<td>When a teacher makes a determination on student progress, or personal growth, based on criteria and/or careful weighing of evidence.</td>
<td>Evaluate how students are doing in a summative manner over a certain period of time in the form of informal feedback, written feedback on student work, or report cards / progress reports / parent teacher conferences.</td>
<td></td>
</tr>
<tr>
<td>Judge</td>
<td>To form an opinion about through careful weighing of evidence or testing of premises, to form an evaluation of</td>
<td>When a teacher forms an opinion about their students’ progress, their own personal growth, or the system of education to which they belong.</td>
<td>Form an opinion about the state of the educational system based on their experience within the system.</td>
<td></td>
</tr>
</tbody>
</table>

I put in pre- and post- tests into Google docs and um I look at all the kids that are way, that start at the pre-test and they were super far behind and then I go to the kids that are proficient, and you know there’s a whole data chart and like it shows me, I love it – a lot of people hate it because it’s like, extra work, but it really lets me see where they went, from here to where. (Lucas, Q13)

So I’m walking around and I’m mentally checking how they’re using the terminology instead of saying, you know, “as the air goes up”, I hope they’re saying “as the air or temperature increases”, looking for the word “increase”, “cloud.” (Adele, Q2)

I know that they are learning the language because they are able to use it in their response and they are able to synthesize what they've learned and explain it to me in their own language so or their own words (Valera, Q2)

… it’s become such an assessment heavy culture, just in general in the United States in education. (Claire, Q4)
<table>
<thead>
<tr>
<th><strong>Integrate</strong></th>
<th>To incorporate or combine together</th>
<th>When a teacher incorporates different sources of data or information to plan their lessons or to assess students’ progress in the content and language of the unit.</th>
<th>Combine content language strategies and second language acquisition strategies in one’s teaching repertoire to promote the academic achievement of emerging bilingual students.</th>
<th>So I’ve had to kind of mold it (academic language) around, so that way I have all of those elements. And hope that when they observe they notice all of that. So it’s integrated, it’s in there, (Adele, Q2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conclude</strong></td>
<td>To bring to an end in a particular way, or to reach a logically necessary end by reasoning</td>
<td>When a teacher reasons through a logical argument in order to come to an end, or to develop an opinion about a new educational reform, a students’ progress, an assessment system, etc.</td>
<td>Make conclusions with regard to student progress to inform future instruction and to serve as the basis for student evaluations.</td>
<td>…when I don't know exactly what's on the assessment, that's not a measurement of my instruction. That's a measurement of my guess. (Dina, Q4)</td>
</tr>
<tr>
<td><strong>Reflect</strong></td>
<td>To think and act in a manner that is driven by active and persistent consideration</td>
<td>When a teacher thinks and acts in a manner that is driven by active and persistent consideration in regard to their teaching and assessment practice.</td>
<td>Consider the practices that a teacher undertook in her lesson, and question if these practices appeared to increase emerging bilingual students’ access to the content based on their assessment practices. If these practices did not increase emerging bilingual students’ access to the content, what was the source of the problem?</td>
<td>Who did well and who didn't and why was it, but they didn't do well or why was it a question [they] bombed, or what did that look like? (Serena, Q13)</td>
</tr>
</tbody>
</table>
**RQ 2. To what extent do teachers of emerging bilingual students use assessment critically?**

**CE component**

<table>
<thead>
<tr>
<th>Concept / Word</th>
<th>General Definition (all in verb form)</th>
<th>Contextualized Definition (Please note: These definitions are not exhaustive. Do not consider these definitions as limiting to other possibilities that may not be mentioned.)</th>
<th>Example in Context with Respect to Assessment with Emerging Bilingual Students</th>
<th>Example from data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propose</td>
<td>To put forth a plan or intention, such as a new idea or concept</td>
<td>When a teacher develops an educational idea or concept that they put forth to their colleagues, school, or broader educational community.</td>
<td>Suggest to one’s principal a plan to incorporate ongoing informal assessment methods, or ongoing formal assessment methods such as the use of a portfolio.</td>
<td>I do think that they need to have assessments, but it shouldn’t be so much. (Lucas, Q4)</td>
</tr>
<tr>
<td>Imagine</td>
<td>To form a mental image of something not present</td>
<td>When a teacher develops a mental image of an idea or a thing that is completely new in order to improve her teaching or assessment practice.</td>
<td>Imagine new assessment methods that can provide valid and reliable inferences of students’ understanding of language and content in a particular content area.</td>
<td>I’m still trying to find ways to bring in literature, you know, cuz I don’t want them to just respond to prompts! (Alma, Q1)</td>
</tr>
<tr>
<td></td>
<td>To modify according to changing circumstances</td>
<td>When a teacher somewhat modifies a lesson, curriculum, or assessment methods or various reasons such as issues of language in content.</td>
<td>Adapt different assessment methods to fit into one’s own classroom. For example, taking Marzano’s exit ticket slip but instead of or in addition to “content” and “effort” self-assessment ratings for the lesson, add</td>
<td>One example was (pause) density is a commonly misunderstood key point. And so when I retaught density, instead of doing the math part, I tried to teach it from a different conceptual method. And so instead of doing mass</td>
</tr>
<tr>
<td>Adapt</td>
<td>Transform</td>
<td>To consider the merits and demerits of and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adapt</td>
<td>Transform</td>
<td>To consider the merits and demerits of and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“language” as a rating as well.</td>
<td>divided by volume, I gave them 2 same size cups and cotton balls, and they put 2 cotton balls in one cup and 16 cotton balls in the other cup. Which one is more packed in? The one that has more stuff packed into it. And then we talked about how cotton balls can be mass, but the size of the cup is volume. So I usually try to teach it in a different conceptual approach, either going from visual to kinesthetic or kinesthetic to auditory or I usually try to switch learning styles is like my approach. (Claire, Q13)</td>
<td>When a teacher completely changes a lesson, curriculum, or assessment method in order to suit their needs for instruction or assessment.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modify the text of the first half of a test’s questions to be linguistically less complex, retain the second half of the test questions in their linguistically complex form. Then, one can compare how students performed on the standard form questions and the parallel, modified questions to more easily pinpoint if the issue is language or content.</td>
<td>So I spent the first weeks, months almost, drilling multiplication into their head. Because in 5th grade they have to know multiplication in order to do fractions, in order to do all the major concepts that we’re doing in here. So if they don’t have that base, they’re not going to succeed. (Lucas, Q13)</td>
<td>When a teacher finds shortcomings or merits with a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critique assessment methods and assessment policies</td>
<td>And then every test I take that is a standardized test, like</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Criticize</strong></td>
<td>judge accordingly, to find fault with</td>
<td>particular reform or concept and shares their opinion with others, either for constructive or destructive purposes.</td>
<td>to identify weaknesses and to brainstorm about how they can be improved.</td>
<td>we have to do benchmark, and we’ll have to grade them coming up, I just pick those things apart. And then I’m relentless when I pick them apart in front of district people. And then I’m afraid that I’m going to get in trouble, cuz I’m like: “this test is awful, you just translated it into Spanish. Direct translation. From English into Spanish. Just like you put it in Google translate and you just printed it out because it doesn’t even make sense! (Valera, Q13)</td>
</tr>
<tr>
<td><strong>Question</strong></td>
<td>To ask or inquire, to subject to analysis</td>
<td>When a teacher asks themselves or others, either reflectively or publicly, a question about concepts or issues that are relevant to their educational or professional environment.</td>
<td>Actively ask questions and seek answers to issues surrounding assessment methods for emerging bilingual students. When a teacher receives a new assessment for a new unit or as a requirement by a school, the teacher questions if it meets the needs of the students and of the teacher in assisting the learning and teaching process for emerging bilingual students in that content area.</td>
<td>When I find that it comes up a bunch of times, and I’ll ask them too, even as a class, I’m like “Why are we all struggling with 14?” And they’ll, you know, they’ll tell me. I’ve never asked about an item and not gotten some insightful information. I, they never say, “Well I don’t know.” They’ll tell you “because we use this word and so this doesn’t make sense” or whatever, you know, they’re really, it’s helpful. (Jamie, Q4)</td>
</tr>
<tr>
<td><strong>Organize</strong></td>
<td>To arrange into a structure by</td>
<td>When a teacher determines how</td>
<td>Organize a group of teachers to write</td>
<td>I am a legislative person for our</td>
</tr>
</tbody>
</table>
| **systematic planning, or to arrange a social event or movement** | **elements fit or function within a particular structure, such as how to arrange lessons in their unit of study. Also, when a teacher arranges a social event or movement to inspire social, professional, or political change.** | **letters to policy makers to ask for more equitable assessment practices for emerging bilingual students.** | **association and so I take that on but I try and take it on as an individual voter involved person. Researcher: Right. Citizen.**  
  
Dina: Yeah. And encourage other teachers to do the same. (Dina, Q4) |
| --- | --- | --- | --- |
| **Collaborate**  
To work jointly with others or together especially in an intellectual endeavor | **When a teacher works together in a partnership with students, parents, colleagues, supervisors, community members, or legislators to achieve a shared educational goal.** | **Work together with fellow teachers to plan assessments for emerging bilingual students that are valid, reliable, align to the standards, and are helpful to the teaching and learning process.** | …we talk about what um you know, what they’re getting, what they’re not getting, we come up with strategies for whole group and small group instruction, and then we come back a week later to see how it went. (Alma, Q1) |
| **Create**  
To make or bring something new into existence | **When a teacher thinks of a new pedagogical approach, educational innovation, or idea in order to improve the teaching or assessment of emerging bilingual students.** | **Create one’s own test questions instead of relying on a text book. Alternatively, create one’s own lessons to explicitly teach the language structures, functions, and vocabulary that will be tested.** | I designed that based on the Common Core standards, and then design a unit that would help them get there (Serena, Q1) |
RQ 3. Which factors influence teachers’ critical thinking about and critical use of assessment for emerging bilingual students?

**SI and external forces**

<table>
<thead>
<tr>
<th>Concept / Word</th>
<th>General Definition (all in verb form)</th>
<th>Contextualized Definition (Please note: These definitions are not exhaustive. Do not consider these definitions as limiting to other possibilities that may not be mentioned.)</th>
<th>Example in Context with Respect to Assessment with Emerging Bilingual Students</th>
<th>Example from the data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Become Familiar</strong></td>
<td>To get acquainted with something new.</td>
<td>When teachers get acquainted with new changes in curriculum, texts, standards, protocol, etc.</td>
<td>Look at the WIDA standards to see the guidelines according to which students will be evaluated</td>
<td>I do a thing called window shopping where I let them get up and walk and look at other people’s answers and then I’m just watching to see “Are they writing?” “Are they still not writing?” (Jamie, Q1)</td>
</tr>
<tr>
<td><strong>Listen</strong></td>
<td>To pay attention and attend to others’ speech</td>
<td>When a teacher pays attention and attends to students’, colleagues’, parents’, and superiors’ speech. Also, when they pay attention to the educational and political speech in the media.</td>
<td>Pay attention to what administrators explain they want to see in a teacher’s classroom during post-observation conferences or during staff meetings</td>
<td>Then I asked my kiddos who I was worried wouldn't access it, &quot;How many stories are in your home?&quot; “One.” &quot;Okay, how many stories are in your home?&quot; (Dina, Q12)</td>
</tr>
<tr>
<td><strong>Use</strong></td>
<td>To employ a tool or instrument in an activity</td>
<td>When a teacher employs a text, a reference, or other materials when engaged in a lesson or when engaged in assessment</td>
<td>To use the assessment materials provided to you, either required or not required by the administration.</td>
<td>I wish I could say it was my idea but it was something that we pulled from the book (Alma, Q4)</td>
</tr>
<tr>
<td><strong>Apply</strong></td>
<td>To use what you know in a different or new situation</td>
<td>When a teacher uses information in a new way, or applies a known procedure to a familiar task. Specifically, to use information from</td>
<td>Use materials provided at professional developments without modifying them to help emerging</td>
<td>I try to align that exit ticket or that spot check directly to the objective so that it’s really concise and really</td>
</tr>
<tr>
<td>Know</td>
<td>To have information, facts, or recognition of procedures.</td>
<td>When a teacher is able to recall or retrieve relevant information from long-term memory, such as to remember the understandings found in previous readings and educational experiences.</td>
<td>Have a repertoire of teaching methods specifically geared towards emerging bilingual students.</td>
<td>And so I know with the ELLs, here’s what I learned, is that I know you should have more questions to account for those, (Valera, Q4)</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Understand</td>
<td>To comprehend conceptual ideas or how facts/procedures are related to each other.</td>
<td>When a teacher constructs meaning from instructional messages including oral, written, and graphic communication from students, parents, colleagues, superiors, etc. OR developing an idea of student understanding based on information (like gathering in FA).</td>
<td>Understand when to use each teaching method in one’s repertoire</td>
<td>If they give it to me and I’m reading it and I’m like, so can you come here for a second? What does this mean, tell me more about this. And if they give me an answer and I’m still pushing I’m like, well what do you mean by that exactly? And finally sometimes they’re just like, “Are we done???” (Valera, Q4)</td>
</tr>
<tr>
<td>Assume</td>
<td>To take as granted or true, to suppose</td>
<td>When a teacher takes an assertion, point, conclusion, or concept as true without questioning its truth.</td>
<td>Take an idea as true since others believe it to be true, and the teacher doesn’t stop to think about it.</td>
<td>… I guess that comes from the actual test itself, if they can read it, although they should be able to read it in 5th grade now, and um I don’t have any students who can’t read English so um they’ll read it (Lucas, Q2)</td>
</tr>
<tr>
<td>To obey and conform to rules or</td>
<td>When a teacher obeys or conforms to the code of</td>
<td>Obey regulations from the school</td>
<td>… our liaison who comes in and tells us</td>
<td></td>
</tr>
<tr>
<td>Comply</td>
<td>To act in accordance with a certain code of conduct, however formal or informal.</td>
<td>conduct and culture of the school at which they work, in addition to complying with the district, state, and federal regulations.</td>
<td>administration, and district and state officials. For example, place emerging bilingual students in classes according to their test score results.</td>
<td>what we should do, recommendations, and we really are forced to assess every concept that we’ve taught (Adele, Q4)</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Copy</td>
<td>To make a copy of; transcribe; reproduce</td>
<td>When a teacher directly reproduces material for their students’ learning or for assessment purposes without modifying it</td>
<td>Copy an end-of-unit test from the back of the book without consideration for whether or not the assessment is an appropriate tool for gathering students’ understanding of the unit.</td>
<td>Sometimes I’ll copy the vocab definition from the vocab packet onto their homework, so they don’t need to refer back to the vocab packet and it’s just there for them. (Claire, Q3)</td>
</tr>
</tbody>
</table>
Appendix E.

Extended teacher narratives.

Adele

Interview

With respect to System I specifically, Adele mostly discussed the activity of ‘complying’, as it represented approximately half of all System I activities. She did not reference the activities ‘becoming familiar,’ ‘listening,’ ‘using,’ ‘applying,’ ‘knowing,’ ‘understanding,’ and ‘assuming’ as System I activities as frequently as the activity of ‘complying.’ She did not mention the activity of ‘copying’ at all. This means that Adele engages in System I thought and action processes mainly through complying with requirements or conditions set in place by the school and other teachers. As a teacher at the end of her first year, it is logical that she would have a lot of activity centered around compliance, since she is learning the culture of the school as well as the requirements set by the administration, the district, and the state.

However, it is still somewhat surprising that the overwhelming majority of her System I activities are based in compliance. Those types of activities overshadow the rest of her activities at this level, causing one to pause and wonder if this is truly proportional to the activities one might expect from a teacher at the end of her first year. Her more extensive reference to compliance-oriented activities might be explained by the context of the “data team” meetings that took place every week in her school. In these meetings, each teacher was required to have an assignment from their students graded, analyzed, and discussed among their grade team, ready for discussion in the data team meeting. There was paperwork that teachers were asked to complete and bring to these weekly meetings as well. Adele also mentioned that she complies with other mandated practices, such as state testing. She noted, “TCAP took over my instruction
for three weeks. I stopped everything curriculum-wise” (Adele, personal communication, May 15, 2014). This made it clear that Adele needed to comply with mandated state testing practices as well as district and school-specific assessment practices.

With regard to System II, Adele mostly discussed the activity of reflecting, as it represented some of the System II activities. She also frequently referenced ideas that represent the activity of ‘concluding’, with it appearing in a few of the System II activities. ‘Judging’ and ‘analyzing’ each appeared in a few of the System II activities. She did not reference ‘comparing,’ ‘contrasting,’ ‘interpreting,’ ‘evaluating,’ or ‘integrating’ as System II activities as frequently as the aforementioned activities, if at all. All System II-type activities were represented in Adele’s interview, however. She engages in System II thought and action processes mainly through reflecting, which seems characteristic for a teacher at the end of her first year. Essentially, when someone is learning a new practice, one would expect that she would give persistent consideration to the effectiveness of techniques used in this new practice and how to shift her activities or frame of thinking accordingly.

Adele readily volunteered to discuss the area where she identified herself as weak. She explained that she knew where she was weak because her LEAP observation had identified her as being “poor” at academic language (Adele, personal communication, May 12, 2014). One could sense a certain amount of frustration in her reflection on this topic regarding whether or not her strategies were “enough” to please her observers. She clearly struggled with academic language in her practice:

I used to have a notebook where they wrote the terms down, but that was also still not enough. So you’ve got to make sure that they are verbally talking with it and they feel comfortable and they know what the term means. ‘Cloud’ is something simple, but you
know, how it functions is completely different. So you have to make sure they understand
that. No pressure. Let me just figure that out, make sure that every 30 kids has said every
word and make sure that they understand, wait, yeah (Adele, personal communication,
May 12, 2014).

Adele communicated that it is important to improve in the area of academic language, but she
indirectly concluded that she did not feel supported enough to achieve these goals. This, in turn,
made her feel quite pressured to perform. She concluded that her educational experience did not
prepare her to teach emerging bilinguals: “I went to school and they didn’t teach us. There was
no class on ELLs. There was a class on differentiation, and that talked more about the fact that
there’s low kids, high kids, medium kids, there are special ed[ucation] kids, and there are some
with learning disabilities. But there was no really big discussion on, I mean, maybe like a chapter
of a book” (Adele, personal communication, May 12, 2014). She reflected on the fact that “I feel
like I haven’t been trained enough in it … so while the learning curve has been very steep for
me, they don’t really give you time or an opportunity to mess it up” (Adele, personal
communication, May 12, 2014). She expressed that she looked forward to the Public School
District’s summer classes to become certified to teach emerging bilinguals because she struggled
with this to such an extent.

Adele also engaged in the activities of ‘analyzing’ and ‘judging’ a fair amount of the time
in the System II domain. These activities might be somewhat explained by the central presence
of the ‘data team’ meetings in her school as well. In these weekly meetings, teachers made
judgments and came to conclusions about their students’ progress through analysis of their
assessment data. In addition, Adele came to conclusions about the effectiveness of these
meetings, and this approach overall, in the interviews I had with her. Adele concluded that she
would “love to have data team maybe a lot more. As obnoxious as it is” (Adele, personal communication, May 15, 2014). To bridge into the CE component discussion, Adele even proposed that given “how much it has showed growth, we need to do it with every subject” (Adele, personal communication, May 15, 2014).

With respect to the CE component, Adele mostly discussed the activity of ‘adapting,’ as it represented some of the CE component activities. She also heavily referenced ideas that represent the activity of ‘collaborating,’ with it appearing in some of the CE component activities. The activities of ‘adapting’ and ‘transforming’ were simultaneously referenced in a few of the CE component activities, as was the activity of ‘proposing.’ The activity of ‘transforming’ on its own, ‘questioning,’ and ‘creating’ were not mentioned as frequently as the aforementioned activities. She did not reference the activities of ‘imagining’ or ‘organizing’ at all. This means that Adele engages in CE component processes mainly through the activity of ‘adapting’. After learning about the context of the data team meetings in her school and knowing that the school has a culture of prioritizing data analysis, it makes sense that Adele focuses on adapting her future lessons based on her analysis of her students’ data every week.

Adele speaks of a transformation in her practice which shows that she did not always focus on adapting her lessons based on students’ data: “at the beginning of the year it was ‘follow the curriculum, follow the curriculum, oh my gosh stay true to what, I hope I’m on the right day when I’m supposed to be’ vs. ‘the curriculum’s great, but what do my kids need?’ … But I still follow it, but I just, it’s not military, ‘let’s go, follow the deadline.’ It’s kind of ‘go with it, take out what you need, cut out, add, whatever you need’” (Adele, personal communication, May 12, 2014). She does this through collaboration with her colleagues, which explains why collaboration is also a focus in her discussion. This collaboration is mandated,
though, as illustrated by Adele’s comments on the topic: “we have been fine with collaborating. Overall, I mean that’s never really been an issue. But it has forced us to collaborate, because we don’t want to show up not ready. Because we don’t wanna deal with the consequences” (Adele, personal communication, May 15, 2014). When asked what the consequences were, she replied in reference to the administration and the district liaison, “Well you know they’ll yell at you. Not yell at you, but lecture you and you don’t wanna be put under a bad tone towards them” (Adele, personal communication, May 15, 2014).

Although this analysis and collaboration is rather forced upon her by the school, she also believes that the strategy works well in her practice, and she hinted that she would probably do it on her own at this point. This signals a bit of a transition for her in terms of autonomy. She noted in her discussion of these meetings: “They [the administration] really won’t cut you slack now. It’s very data – data – data, intense. So you can’t, you have to show up to the meeting, and you have to be ready and you have to be ready to have a great conversation, cuz … if they don’t, then you know. Trust me. We know!” (Adele, personal communication, May 12, 2014). The first part of this quote focuses on her compliance in preparing for these meetings. The last part of this statement reveals an interesting twist on the level of ownership she has with regard to these meetings. She suddenly exclaims that “you know” and “we know” when she talks about how other teachers may come unprepared to these meetings. The shift to expressing this sentiment in the “we” form shows that she holds others accountable to be prepared for these meetings, while she herself is still also accountable to the administration for being prepared. Thus, there is a palpable tension between her ownership of adapting her instruction based on data analysis, while the motivation for this activity still belongs somewhat to the mandates and the culture of the school.
Adele also discussed ideas that represent the CE component-type activity of ‘proposing’ when she was sharing her opinion on the data team meetings held at her school. She proposed that, since she thinks they are beneficial and have shown growth in student performance, “we need to do it with every subject” (Adele, personal communication, May 15, 2014). She proposed having a set time every day to engage in this type of meeting, with a different subject as the focus for each day. She feared that other teachers would dislike her if they heard her propose this idea, yet she came up with ways in which her colleagues might begin to like the meetings more. She thought it could be less taxing on teachers to prepare for these meetings if, “instead of reading the [artifacts for the] whole class, grab six people’s, and you would know who’s kind of high, low, medium, and then we could totally do it for every subject” (Adele, personal communication, May 15, 2014). She shared this proposal with me excitedly as she thought of the ways they could structure data team meetings as a more frequently occurring part of their week. Adele even proposed ideas about the formal, large state assessment program. She explained that she thought that the Transitional Colorado Assessment Program should be “rewritten so it’s not like, ‘I gotcha’ or ‘I fooled ya,’ it’s more like, ‘What do you know?’ and ‘Tell us all the things you know’ (Adele, personal communication, May 15, 2014). She thought that there were tricky questions that were misleading to students, and that the test could benefit from being more straightforward in nature. Adele used CE component thinking to propose several interesting ideas over the course of the interviews.

To summarize, Adele engages in critical thought and action frequently through reflecting, concluding, judging, analyzing, adapting, collaborating, and proposing. Her activities reside mostly within the sphere of System II. One would categorize her self-identified behavior, then, as residing in the middle on the spectrum of critical thought and action with regard to assessment
for emerging bilingual students and beyond. Given the way in which she expresses increasing levels of ownership of some required practices, she may push herself higher on the spectrum of critical thought and action over time in the profession.

**Observations**

With respect to System I specifically, Adele mostly exhibited activity that represented ‘listening,’ as it comprised some of her System I activities. She also displayed activity that represented ‘becoming familiar,’ with it appearing in some of the System I activities. She displayed activity that represented ‘using,’ as it was found in a small percentage of the System I activities. She did not exhibit ‘knowing,’ ‘understanding,’ ‘applying,’ ‘complying,’ ‘assuming,’ or ‘copying,’ as System I activities as frequently as the aforementioned activities, if at all. This means that, from the classes observed, Adele engages in System I thought and action processes mainly through becoming familiar with what students understand with respect to the content objectives, by listening to students construct knowledge during class, and by using various resources for her teaching. She became familiar with her students’ understanding of content by administering a formal pre-test to her students on the topic of poetry (Adele, field notes, May 13, 2014). Although a poem written in five minutes may not furnish necessarily useful information for Adele, the pre-test theoretically gave her some information about what her students understood with respect to poetry. She did listen to her students extensively to learn more about their thinking process. She asked many questions of students and listened to their responses. When students worked on defining the idiom “burying the hatchet,” she listened to many students’ responses before giving her own example (Adele, field notes, May 13, 2014). She also used many resources to support her instructional objectives. She made use of videos to help
provide examples of figurative language for students, for example. She also used computers with her students so they could type their final research papers.

With respect to System II specifically, Adele mostly exhibited activity that represented ‘integrating,’ as it comprised half of all System II activities. She displayed activity that represented ‘judging,’ with it appearing in the other half of all System II activities. She did not exhibit ‘comparing,’ ‘contrasting,’ analyzing,’ ‘interpreting,’ ‘evaluating,’ ‘reflecting,’ or ‘concluding’ as System II activities at all. This means that, from the classes observed, Adele engages in System II thought and action processes mainly through integrating her lessons with real life experiences and by judging the quality of students’ responses. She integrated real life experiences into her lessons when she chose to teach a song that students were familiar with in order to help teach poetry. She used the lyrics of the song to help illustrate some literary devices. Additionally, Adele judged students’ responses by providing affirmative or negative feedback when students provided their thinking to her in class.

With respect to the CE component specifically, Adele only exhibited activity that represented ‘adapting’ and ‘transforming,’ as this pair of activities comprised all of the CE component activities. She did not exhibit ‘proposing,’ ‘imagining,’ ‘collaborating,’ ‘organizing,’ ‘questioning,’ ‘criticizing,’ or ‘creating’ as CE component-related activities at all. This means that, from the classes observed, Adele engages in CE component-related activities through adapting and transforming her feedback to students based on their responses and reactions in class. For example, she gave corrective linguistic feedback to her emerging bilingual students. She told one student who used the word ‘effective’ in his research paper, “You and the word ‘effective’. Effect can affect you, but effective is different. Like, ‘I put hair spray in my hair because it’s effective at keeping it in place.’ Is effective what you meant?” (Adele, field notes,
May 12, 2014). She modified the way she instructed students based on the language they used in their writing.

To summarize, Adele engages in critical thought and action relatively infrequently in observed classroom situations. Therefore, one would categorize her as residing lower on the spectrum of critical thought and action with regard to observed activity in the area of assessment for emerging bilingual students.

**Alma**

*Interview*

With respect to System I specifically, Alma mostly discussed the activities of ‘using’ and ‘applying’ simultaneously, as these paired activities represented some of her System I related activities. She also heavily referenced ideas that represent the activity of ‘complying’, with it appearing in some of the System I activities. ‘Becoming familiar’ was a frequently represented activity by Alma, also comprising some of the System I activities. She did not reference ‘listening’, ‘knowing’, ‘understanding’, ‘copying’, or ‘assuming’ as System I activities as frequently as the aforementioned activities, if at all. This means that Alma engages in System I activities mainly through using and applying new assessment-related strategies for her emerging bilingual students. Many times Alma would refer to using new resources and applying new strategies in order to meet the demands of new assessment initiatives. For example, she used Transitional Colorado Assessment Program resources to see how test items might appear for her students. She incorporated district-approved, nationally-distributed assessments in order to monitor students’ progress in each subject. She even used a book as a resource to initiate self-reflections for students’ writing. Alma noted that “the idea is that data drives our instruction and we’re using the book to help guide us” (Alma, personal communication, April 18, 2014). From
this discussion and from learning about the data-team meetings, it is apparent that data-driven instruction is a central goal of the district. To reach this goal, teachers use all resources to which they have access.

Alma shared ideas that represent the activity of ‘complying’ with regulations of her school when she discussed the limitations of the curriculum at her school. The teacher noted that she had enough experience to fit in teaching concepts that she found to be important, such as how to work towards social justice or how to think critically. She expressed concern for teachers who have less experience than her because there are “new teachers that are, are not, being mentored to do that in that way, or [who are] feeling like there’s not a moment to spare. You know, ‘I don’t have time to do a ten minute read aloud and talk about Sounder,’” (Alma, personal communication, April 18, 2014). She acknowledged the restrictions on what could be taught indirectly by explaining how there is not enough time to teach the concepts she thinks are critical for students.

Alma also shared ideas that represent the activity of ‘becoming familiar,’ for example, when she talked about how she familiarizes herself with her emerging bilinguals. She discussed how one of her newest emerging bilinguals in class has struggled because the language can be overwhelming at times, but that she thinks she is very capable of overcoming this challenge. She explained, “I can see that she’s, I know how to do this, I just need to be taught the language around it, and I need to be taught the concept, but I know how to do the math. I know how to do the math!” (Alma, personal communication, April 18, 2014). Alma became excited at the promise of some of her emerging bilinguals’ abilities.

With regard to System II, Alma mostly discussed the activity of ‘reflecting,’ as it represented approximately half of all System II activities. She also referenced ideas that
represent the activity of ‘concluding,’ with it appearing in a few of the System II activities. She did not reference ‘comparing,’ ‘contrasting,’ ‘analyzing,’ ‘interpreting,’ ‘evaluating,’ ‘judging,’ or ‘integrating,’ as System II activities as frequently as the aforementioned activities, if at all. This means that Alma engages in System II activities mainly through reflecting on various aspects of her practice as a teacher. She reflected at length on how to make her teaching social justice-oriented despite initiatives that force most teachers to focus their time on data-driven instructional practices:

I think because of my experience I can still sneak in reading about the Holocaust and …. really talk in depth about those social justice kind of things so that education is still transformative in nature. So it’s not just, ‘I’m gonna take this test’. It’s really what makes a school relevant, you know, why do I need to know this, you know? (Alma personal communication, April 18, 2014)

Alma also spent a considerable amount of time reflecting on her relationship with the data-driven initiative taking place at her school. Specifically with regard to language, she reflected on how to incorporate opportunities during classwork for her emerging bilingual students to use language authentically. She thought critically about her own daily practice, noting “where I can probably structure that conversation more is when they move to work with their partners” (Alma, personal communication, April 18, 2014). Finally, Alma reflected on what motivates her to be a teacher by comparing the experience of emerging bilinguals in her class to how she felt as an emerging bilingual student in grade school. She explained that “they feel dumb and no one’s there telling them that they’re not” (Alma, personal communication, April 18, 2014). Alma confided to me that she had the same poor self-concept until she was a senior in high school, at which point
someone believed in her capabilities. Alma routinely reflects on her own experience to empower
her emerging bilingual students through her pedagogical practice.

Alma also shared ideas that represent the activity of ‘concluding’ when she discussed
how teachers need to focus on students’ socio-emotional needs in addition to their academic
needs. She concluded that you can’t measure a student’s socio-emotional well-being. However,
she posited that she thinks it is still very important to address because “that affective filter can
really get in the way of stuff and so I feel like if they feel, I feel like if this sense of safety and
security and I know no one’s going to laugh at me,” then it can help them to focus on their
lessons (Alma, personal communication, April 18, 2014). Alma was able to conclude that socio-
emotional well-being is important and therefore she prioritized students’ socio-emotional
development in her class.

With respect to the CE component, Alma mostly discussed the activity of ‘adapting’ and
‘transforming’ simultaneously, as these paired activities represented a few of the CE component
activities. In addition, Alma’s activities represented ‘adapting’ on their own, as they constituted a
few of the CE component activities. She also referenced ideas that represent the activities of
‘questioning’ and ‘criticizing,’ with both appearing separately in a few of the CE component
activities. She did not reference ‘proposing,’ ‘imagining,’ or ‘organizing’ as CE component
activities frequently, if at all. She referenced ‘collaborating’ and ‘creating’ a small percentage of
the time (12% each) as CE component activities. This means that Alma engages in CE
component activities mainly through adapting and transforming together to change her lessons
based on the needs of her students. She described how she does this both in an academic and in a
socio-emotional context. In an academic context, Alma described how she examines district
resources for the current content, and if it’s not “the exact skill I’m honing in on that day, I will
modify it or I won’t use it or I’ll postpone using it until I know they’re ready for it” (Alma, personal communication, April 18, 2014). In this way, she adapts her class activities and transforms classroom resources to fit the needs of her students on that particular day. In a socio-emotional context, Alma discussed how she transformed the first few minutes of the day in her classroom to adapt to the socio-emotional needs of students. She tells students the expectations and planned activities for the day because they deserve to know that and to be able to take the time to absorb that information at the beginning of the day. She also allows a few additional minutes for students to unpack so that she can check in with students to see “OK who looks like they’ve had a rough night?” because “they have to be learning ready” (Alma, personal communication, April 18, 2014). Alma has found that she needs to adapt to both the content-based and affective-based conditions of her classroom in order to successfully help her students progress in the school environment.

Alma also shared ideas that represent the activity of ‘questioning’ when she discussed the recent data-driven emphasis at her school. She wondered, “I’ve asked quite a bit, understanding, I’m a seasoned teacher, and if I’m feeling a little bit lost and asking for support, what are my colleagues feeling, knowing that they’re brand new and aren’t really used to having to come up with their own stuff?” (Alma, personal communication, April 18, 2014). She was concerned that the new teachers would not be able to progress with the little support they had been given in order to make the mandated changes to their practice. She also discussed ideas that represent the activity of ‘criticizing’ when she expressed discontent that professional development has not differentiated for teachers of various experience and ability levels. She even critiqued:
Alma: I feel like we’re asking our teachers to differentiate everything for our kids, but we’re not differentiating how we’re asking, you know, like our delivery of this professional development, in essence, is just, [pause]

Researcher: One size fits all?

Alma: Yeah. But we’re not gonna tell you what size it is! (Alma, personal communication, April 18, 2014)

Alma both questioned and critiqued the way in which professional development was being carried out, even while she hesitantly embraced the content delivered in the professional development.

To summarize, Alma engages in critical thought and action quite frequently by adapting and transforming her instruction and assessment based on the needs of her emerging bilingual learners. She also considers the needs of students in socio-emotional areas as well as in the academic realm. It is of utmost importance for teachers to critically assess socio-emotional aspects of learning for emerging bilingual students, as Alma does in her practice. Alma even engages in these activities despite the fact that her school does not promote or even necessarily approve of these activities. Therefore, one would categorize her self-identified behavior as residing high on the spectrum of critical thought and action with regard to assessment for emerging bilingual students and beyond.

Observations

With respect to System I specifically, Alma mostly exhibited activity that represented ‘listening,’ as it comprised approximately half of all System I activities. She displayed activity that represented ‘becoming familiar,’ with it appearing in some of the System I activities. She also displayed activity that represented ‘using,’ as it was visible in some of the System I
activities as well. She did not exhibit ‘knowing,’ ‘understanding,’ ‘applying,’ ‘complying,’
‘assuming,’ or ‘copying,’ as System I activities as frequently as the aforementioned activities, if
at all. This means that, from the classes observed, Alma engages in System I activities mainly
through listening to her students, becoming familiar with their needs, and using various resources
in her instruction.

Alma listened to her students when she elicited their thoughts during whole group
instruction and when she circulated around the room during students’ work time. She was not explicit in reiterating what she heard, yet based on her feedback to students, it is clear that she listened to students both when she questioned them directly and when she overheard them as they worked. She became familiar with students’ needs when she collected information about students’ understanding. She did this by asking students to raise their hands if they understood a concept. For example, she asked students if they could recall the difference between area and perimeter (Alma, field notes, April 18, 2014). She would also become familiar with students’ understanding by asking students to complete an exit ticket, or small quiz, before the end of class. One of exit tickets covered converting percentages to decimals, for instance. Alma also used resources to help support her instruction. For example, Alma used small whiteboards for students to practice math problems individually (Alma, field notes, April 18, 2014). This allowed for students to receive instant feedback on their answers as well. Alma also used videos to provide another modality for students to access the content area objective (Alma, field notes, April 28, 2014). These resources served to give students different methods to use to make progress towards the learning goals.

With respect to System II specifically, Alma mostly exhibited activity that represented ‘integrating,’ as it comprised half of all System II activities. She also displayed activity that
represented ‘judging,’ with it appearing in the other half of all System II activities. She did not exhibit ‘comparing,’ ‘contrasting,’ analyzing,’ ‘interpreting,’ ‘evaluating,’ ‘reflecting,’ or ‘concluding’ as System II activities as frequently as the aforementioned activities, if at all. This means that, from the classes observed, Alma engages in System II activities mainly through integrating real life experiences into her lessons and through judging students’ responses through corrective feedback.

She integrates real life experiences into her lessons on a regular basis. She asked her class, when she began a lesson on how to write a book recommendation, “So why on earth would we need a book recommendation?” (Alma, field notes, April 28, 2014). She then pursued a conversation with her class about the relevance of book recommendations to their lives. This served as an introduction to the topic and as a motivating factor for students to learn about book recommendations. She also integrated students’ cultural backgrounds into her instructional practice. During a guided reading group, Alma brought in a discussion of students’ cultural background when they read about students who had an assignment to bring a meal to class that represents their cultural heritage. Alma asked the group about their own cultural heritage. She inquired, “If you had had this assignment, what would you have brought?” (Alma, field notes, April 28, 2014). One student described Spanish rice and soup with “that pink stuff,” to which the teacher responded, “pozole?” (Alma, field notes, April 28, 2014). The student responded excitedly that the teacher was able to recognize the name of the type of soup she eats at home.

Alma also judges students’ responses through explicit feedback. She specifically provided linguistic feedback during the second observation, when the teachers’ language-oriented goal was for students to learn to use the verb ‘convert’ instead of the verb ‘did’ when they described the process of converting percentages to decimals. By asking students to describe
the process of conversion, she also asked them to perform the language function of ‘describing.’ She corrected a single student twice with a quick judgment through feedback. By the third time the student described the process, he used the word ‘convert’ instead of ‘did.’

With respect to the CE component specifically, Alma only exhibited activity that represented ‘adapting’ and ‘transforming,’ as this pair of activities comprised all CE component activities. She did not exhibit ‘proposing,’ ‘imagining,’ ‘collaborating,’ ‘organizing,’ ‘questioning,’ ‘criticizing,’ or ‘creating’ as CE component activities. This means that, from the classes observed, Alma engages in CE component activities through adapting and transforming her instruction to meet the needs of her students.

She adapted and transformed her instruction when, for example, she noticed that students did not pronounce the “ths” at the end of the word “hundredths” (Alma, field notes, April 28, 2014). Alma then gave a mini lesson on plural nouns to help students understand the importance of pronouncing this ending. She explained how adding an ‘s’ to the word “boy” to make it “boys” changes the meaning of the word. Although the word ‘hundredths’ is not plural, she related the plural nouns to the fraction to express how a small difference in the ending of a word can change the meaning of the word. This instructional point was unplanned and was simply a response to the fact that students were not pronouncing the ending of the word “hundredths” (Alma, field notes, April 28, 2014). This teacher also adapted her instruction by providing native language support in Spanish to a student who had missed the prior week of instruction. She used native language to scaffold the lesson for this student in order to help her catch up from the instruction she had missed the week before (Alma, field notes, April 18, 2014). Additionally, Alma modified her instruction for a student who was not able to solve the division problem: “30 / 2” (Alma, field notes, April 18, 2014). The teacher noticed the student struggling to solve the
problem, so she drew two bubbles on the student’s paper and drew 30 dots in each of the two bubbles. The student was able to solve the problem with the help of this modified instruction, and Alma would not have done this had she not first seen that the student was encountering difficulty in solving the problem.

To summarize, Alma engages in critical thought and action with some frequency by adapting and transforming her instruction in observed classroom situations. Therefore, one would categorize her as residing in the middle of the spectrum of critical thought and action with regard to observed activity in the area of assessment for emerging bilingual students.

**Valera**

*Interview*

With respect to System I specifically, Valera mostly discussed the activity of ‘complying,’ as it represented approximately half of all System I activities present in her interview. She referenced ideas that represent the activity of ‘understanding,’ with it appearing in a few of the System I activities. She also referenced ‘knowing’ in a few of the System I activities. She did not reference ‘becoming familiar,’ ‘listening,’ ‘using,’ ‘applying,’ ‘assuming,’ or ‘copying,’ as System I activities as frequently as the aforementioned activities, if at all. This means that Valera engages in System I activities mainly through complying with regulations set by the school administration or school culture.

The most poignant example of compliance that emerged in the interview with Valera involved the grading process. Valera explained that the grades they gave to students on report cards must be consistent with the student’s standardized test score, as requested by her administration. Valera struggled with this requirement because “a lot of times, even if they show proficiency in a task here in the classroom, because I don’t have proof of a standardized test like
our benchmark test, I cannot give them a proficient grade” (Valera, personal communication, March 28, 2014). This demonstrates that Valera’s grading system was not as valued by the administration since it did not “count” unless it matched the standardized test scores given to the student. Valera may well be a critical user of assessment for her emerging bilingual students, yet if her administration only trusts the scores from a standardized test report, ultimately the standardized test score is the information that is reported to students, parents, and the students’ future middle schools. The teacher’s assessment of her students’ performance is disregarded. Granted, the standardized test scores and the teacher’s assessment of her students’ performance should correspond to one another. However, there is always error in testing, particularly when one is testing students on their content understanding in English as they are learning the English language (Solano-Flores, 2008).

Additionally, the requirement to use the standardized test score on a student’s report card undermines the teacher’s professional assessment and judgment. This requirement changed the way in which Valera uses assessment in her classroom. She explained that she uses “assessments to guide my teaching over anything else because, unfortunately, the message that sends to me is that my assessments don't really matter” (Valera, personal communication, March 28, 2014). This practice also affected the students in Valera’s classroom, as they were in fifth grade at the time of the interview and were applying for entrance to middle schools. Their district reviews students’ report cards for entrance into the more highly regarded middle schools, “and so when their report cards don’t necessarily reflect what they’re capable of doing, that makes it difficult for them to get into good schools” (Valera, personal communication, March 28, 2014).

Valera also raised issues related to complying with regulations when she described her interactions with colleagues and the administration about best practices for emerging bilingual
students. Valera had reflected extensively about best practices for her emerging bilinguals due to her educational experience in a CLD program at a local university. However, she felt pressure to keep her new knowledge and understandings to herself because of the culture of the school to conform to practices that follow the status quo. She expressed that “there’s a couple of teachers who get it, but aren’t going to say anything… I guess she doesn’t want to ruffle anybody’s feathers” (Valera, personal communication, April 9, 2014).

Valera expressed ideas that corresponded with the activity of ‘understanding’ when she discussed why she began adding an English Language Development (ELD) block to her instructional day. She explained that she was told she was to be teaching an ELD block when she began teaching, but that no other teachers really carried it out, so she didn’t make it a priority either. Then, she recounted that she took a class with a professor at a local university, and she finally realized, “I’m supposed to be teaching English Language Development!” (Valera, personal communication, April 9, 2014). This is precisely the kind of understanding that lays the groundwork for teachers to use assessment critically for their emerging bilingual learners.

Valera also discussed ideas that represented ‘knowing’ when she shared about what she knows about language supports for emerging bilingual students from professional development sessions. Valera, like Alma, also criticized the way in which professional development for teachers is not differentiated according to the needs of various teachers. Valera explained that she knows many of the themes that are the focus of professional development sessions at her school. For example, Valera claimed that their professional development focused on the classification of Tier One, Two, and Three vocabulary words. She said, “I have the book at home because I use this book when I’m writing lesson plans” (Valera, personal communication, April 9, 2014). For this reason, Valera knows about these types of tiered vocabulary words quite well.
This knowledge also serves as a good building block for understanding how to teach emerging bilingual students.

With regard to System II, Valera mostly discussed the activity of ‘reflecting,’ as it represented some of the System II activities. She referenced ideas that represent the activity of ‘concluding,’ with it appearing in a few of the System II activities. She also referenced ‘interpreting’ in a few of the System II activities. She did not reference ‘comparing,’ ‘contrasting,’ ‘analyzing,’ ‘evaluating,’ ‘judging,’ and ‘integrating’ as System II activities as frequently as the aforementioned activities, if at all. This means that Valera engages in System II activities mainly through reflecting and concluding. When she reflected in her interview, one theme that Valera touched on was how her assessment practice has changed over time as a result of her CLD coursework at a local university and as a result of teaching experience. She discussed how “before it had a negative impact on my psyche, but now I see it more as a tool that I can use to inform my teaching practices” (Valera, personal communication, March 28, 2014). She described a transition from viewing assessment as a failure on her part to now finding where gaps are and filling them, like “checking the oil in your car” (Valera, personal communication, March 28, 2014). Her reasoning for the change in her view on assessment was attributed to the idea that her lessons are now centered on her students, whereas before they were centered on her. She admitted that “my lessons used to reflect my knowledge and what I was capable of doing instead of being written on the needs of my students” (Valera, personal communication, March 28, 2014). It was clear that Valera had reflected on changes in her perspective and practice prior to the interview, as she was readily able to share her reflections with me during the interview. She also referenced topics from her current, habitual reflection on her teaching practice. She noted that she read something in class about the concept of false negatives that really bothered
her. She expressed deep concern for students who “say they got it when they actually didn’t. Because what if they didn’t get it, and … And I move on? That safety net is gone” (Valera, personal communication, April 9, 2014). Valera went on to describe activities that she carries out in her regular assessment practice to tackle this issue, all involving a heightened scrutiny in the way in which test items are worded and structured. She pays particular attention to language in written assessments for her emerging bilinguals. These activities stem from her continued reflection on how to improve her assessment practice for emerging bilinguals.

With regard to Valera’s reference to “concluding” as a component-related activity, she discussed how she has made decisions based on analysis or observations she made in the classroom. For example, she shared that she does not ever give different, modified tests to students based upon their language status. She concluded that she did not want them to “feel separate. Because they’ll be like, why are they getting a different test? And they know who the native Spanish speakers are, and they’ll be like, oh, they’re getting the wrong test because they don’t know English” (Valera, personal communication, April 9, 2014). Through this commentary, it is apparent that Valera concluded that providing language-modified tests is not good practice due to the social repercussions of giving different tests to students. Instead, she adapts the assessments she gives to all students such that emerging bilinguals can access the language used in the assessment.

Valera also discussed ideas that correspond to the activity of ‘interpreting’ when she shared how she interprets the scores and results from the state standardized testing program. She explained that she does not interpret the standardized test scores as a direct reflection of what her students can do because “the writing assessment heavily relies on their ability to read on grade level” (Valera, personal communication, April 14, 2014). She interprets results from the
standardized test scores as a reflection of what her students can do when they can access the story. When her students cannot access the stories in the test booklets, it is difficult for Valera to interpret her students’ performance from those assessments. In those cases, she uses other sources of information to better understand their progress.

With respect to the CE component, Valera mostly discussed the activity of ‘adapting’ and transforming’ as a pair of critical activities, as this combination of activities represented some of the CE component activities. She also referenced ideas that represent the activity of ‘criticizing,’ with it appearing in a few of the CE component activities. She did not reference ‘proposing,’ ‘imagining,’ ‘questioning,’ ‘organizing,’ ‘collaborating,’ or ‘creating’ as CE component activities as frequently as the aforementioned activities, if at all. This means that Valera engages in CE component activities mainly through adapting and transforming her assessments as well as through criticizing her district and school culture with regard to emerging bilingual students.

Valera shared how she creates and adapts her assessments to the needs of her emerging bilinguals, sometimes to such an extent that she becomes frustrated with the process: “I write the same question 100 times and it drives me nuts. Trying to get it right! And I know with the ELLs, here’s what I learned, is that I know you should have more questions to account for those [language-related uncertainties]” (Valera, personal communication, April 9, 2014). Valera pursued further knowledge of her students’ language and content understandings when she explained how she asks them to come to her and verbally explain their answers to questions in addition to their written answers so that she can better ascertain their intended meaning. She expressed concern that she was not getting the right kind of data from the assessments she creates, despite her efforts.
In addition to increasing the quantity of items and clarifying students’ responses verbally, Valera adapts her assessments specifically with regard to the language used in each item. Valera shared that she typically creates two items that target identical content knowledge. The difference between the two items is that “the vocabulary will be different, so I know that my native English speakers might get [item] three, my non-native English speakers might not get [item] three, but they’ll get [item] 13 because they understand the language of it” (Valera, personal communication, April 9, 2014). In this way, Valera can better deduce if students are struggling with the content knowledge or if they are struggling with the language through which the content knowledge is communicated.

Valera also adapts her assessment practice with emerging bilinguals on a daily basis in the classroom. She discussed how she asks her students to complete exit slips after most lessons so “I know if it was me or if it was the content itself” and can adapt her subsequent lessons accordingly (Valera, personal communication, April 9, 2014). She believes that if a couple of students didn’t understand the content, then it should be retaught in a different manner so they can access it. She distinguished this situation from instances when “I have the majority of my students that didn’t get it, then that was clearly on me” (Valera, personal communication, April 9, 2014). She discussed how she might weave the content into the subsequent discussion so that she can re-introduce the content that students did not previously master.

A large portion of Valera’s CE component-type activities involved ‘criticizing’. This was sometimes due to the events that transpired when Valera learned extensively about best practices for emerging bilinguals during her CLD educational experience at a local university, and then spoke to her colleagues and administration about these issues. Valera decided that “graduate school has been the best thing for me and it’s been the worst thing for me because I get into a lot
more trouble” (Valera, personal communication, April 9, 2014). Valera discussed how she critiqued the standardized tests chosen by the district to district personnel when they are poorly translated, for example, in the hopes the district would improve their assessment products and procedures. However, Valera did not feel as though her colleagues supported her in this endeavor, so she also criticized her colleagues for their lack of awareness of EECD issues. She noted that “when you’re in a staff and only I think two of us have an EECD background, in a bilingual school, nobody else is behind that… They just think you’re being difficult” (Valera, personal communication, April 9, 2014). She repeatedly criticized, in a disappointed tone, the lack of support for emerging bilinguals in both her school and in her district.

To summarize, Valera engages in critical thought and action frequently by adapting and transforming her assessment practice as well as through criticizing the issues regarding assessment of emerging bilinguals in her school and district. She not only criticized these issues privately, but she also spoke up to district personnel and the administration at her school about these issues. In fact, she was so outspoken about these issues that she even admitted that it has gotten her into “trouble” (Valera, personal communication, April 9, 2014). She left her position at her school that academic year so that she could pursue a position as an English Language Development coach where she would be able to have more power to ensure that the emerging bilinguals in her jurisdiction are taught and assessed using best practices according to the research base on English language acquisition. She actively seeks opportunities to improve assessment practices for emerging bilinguals even when it is potentially damaging to her own livelihood. These are instances where CE component-type activities are most difficult to engage in, since there is a fear that doing what one has critically determined to be ‘right’ or equitable will threaten one’s own well-being. Therefore, one would categorize her self-identified behavior
as residing very high on the spectrum of critical thought and action with regard to assessment for emerging bilingual students and beyond.

Observations

With respect to System I specifically, Valera mostly exhibited activity that represented ‘listening,’ as it comprised approximately half of all System I activities. She also displayed activity that represented ‘becoming familiar,’ with it appearing in some of the System I activities. She did not exhibit ‘knowing,’ ‘understanding,’ ‘using,’ ‘applying,’ ‘complying,’ ‘assuming,’ or ‘copying,’ as System I activities as frequently as the aforementioned activities, if at all. This means that, from the classes observed, Valera engages in System I activities mainly through listening to her students and becoming familiar with their learning needs.

Valera listened to her students when, for instance, she conducted a running record of a student’s reading skill using leveled readers. She listened very intently and took notes of all of the student’s miscues. She also listened carefully to students as they conferenced with her in peer writing groups, as she gave precise feedback to students in these groups. Valera became familiar with students’ needs when she heard students share their writing responses to a prompt with the whole group. She noticed that some students did not respond to the prompt and that they discussed another topic in their writing instead. Valera followed up by asking everyone to raise their hands if they also did not address the prompt. By doing this, she was able to determine if the whole class struggled in addressing the prompt. She elicited a lot of information during her classroom discussions and through her exit slips, which allowed her to become familiar with her students’ understanding of the content as well.

With respect to System II specifically, Valera exhibited activity that represented ‘comparing,’ as it comprised half of all System II related activities. She also displayed activity
that represented ‘evaluating,’ with it appearing in the other half of all System II activities. She did not exhibit ‘contrasting,’ analyzing,’ ‘interpreting,’ ‘judging,’ ‘integrating,’ ‘reflecting,’ or ‘concluding’ as System II activities at all. This means that, from the classes observed, Valera engages in System II activities through comparing vocabulary and through evaluating a moral issue with students. Valera compared vocabulary words in English to their Spanish equivalent with students in order to help them make cross-linguistic connections. She explained to her students, “A serial is like a story that keeps going, like a novela. It’s a story that keeps going and going and going” (Valera, field notes, April 14, 2014). She compared a novela to a serial story in order to help students access the significance of a story that is defined as a ‘serial.’ Additionally, she evaluated a moral issue with students regarding whether it is ever acceptable to lie to someone else. She facilitated the discussion, but allowed students to interact with each other relatively freely. Students volunteered their opinions in the format of a community meeting rather than in a teacher-led, question and answer format. Thus, the entire class was mutually engaged in evaluating whether it is ever acceptable to lie to someone else.

With respect to the CE component specifically, Valera mostly exhibited activity that represented ‘adapting’ and ‘transforming,’ as this pair of activities comprised approximately half of all CE component related activities. She displayed activity that represented ‘questioning,’ as it was visible in some of the CE component related activities. She also displayed activity that represented ‘collaborating,’ with it appearing in a small percentage of all CE component related activities. She did not exhibit ‘proposing,’ ‘imagining,’ ‘organizing,’ ‘criticizing,’ or ‘creating’ as CE component activities as frequently as the aforementioned activities, if at all. This means that, from the classes observed, Valera engages in CE component related activities mainly through adapting and transforming her lessons based on her students’ needs, questioning students
to elicit meta-cognitive information about their learning, and collaborating with her students by actively engaging students in the assessment process.

Valera adapted and transformed her lessons when she noticed, for example, that students had not adequately addressed a writing prompt. When she realized that many students had not responded in the way she anticipated, she asked students to rewrite their responses. She provided feedback about how to rewrite their responses such that it would address the prompt properly. Valera asked students meta-cognitive questions after she administered a formal, end-of-unit assessment to the class. She asked students in a whole group format if the language of the test was the reason that they answered any items incorrectly. She also asked students to explain their thinking if they answered an item on the test incorrectly. Some students explained that the vocabulary in general was challenging, and others mentioned that they confused two key words: “antagonist” and “protagonist” (Valera, field notes, April 21, 2014). This helped students learn to monitor their own learning better, and it helped Valera broaden her understanding of students’ thinking about the assessed concepts. Valera also collaborated with students by actively engaging them in the assessment process when she asked students to give feedback to each other about their responses to the writing prompt (Valera, field notes, April 14 and 16, 2014). This strategy successfully pushed students to join a community of learners, where everyone was partially responsible for improving the writing by students in the group. Peer assessment was also present during small group work when students worked on their writing response about the book, *Esio Trot* (Valera, field notes, April 14 and 16, 2014). Valera provided feedback to students as well. However, this writer’s workshop especially benefited from the element of peer feedback. Peer feedback made Valera’s classmates collaborate with Valera and with each other as they worked towards the goal of improving each other’s writing.
To summarize, Valera engages in critical thought and action rather frequently during observed classroom situations by adapting and transforming her instruction based on students’ needs, questioning students about meta-cognition, and collaborating with them in the process of assessment. Therefore, one would categorize her as residing relatively high on the spectrum of critical thought and action with regard to observed activity in the area of assessment for emerging bilingual students.

Lucas

Interview

With respect to System I specifically, Lucas mostly discussed the activity of ‘complying’, as it represented some of the System I activities. He heavily referenced ideas that represent the activities of ‘using’ and ‘applying,’ with them simultaneously appearing in a few of the System I activities. He also referenced ideas that represent the activity of ‘knowing’ and ‘understanding,’ with them simultaneously appearing in a few of the System I activities. He did not reference ‘listening,’ ‘becoming familiar,’ ‘copying,’ or ‘assuming’ as System I activities as frequently as the aforementioned activities, if at all. This means that Lucas engages in System I activities mainly through complying with regulations in his school.

For example, Lucas discussed how the administration places a lot of weight on the subjects of Reading, Writing, and Mathematics, which leaves little time for Science and Social Studies. Lucas clarified that “it’s always been Reading, Writing, Math, and it’s never been Science and Social Studies in there” (Lucas, personal communication, April 30, 2014). Lucas expressed some anxiety over how to comply with upcoming regulations regarding Science and Social Studies testing. He explained that “now Science and Social Studies is gonna be tested and it’s gonna be accounted for, so we’re gonna have to make time for that. Big time” (Lucas,
personal communication, April 30, 2014). Lucas had confided in other teachers about how to handle this and it seemed that his colleagues were also concerned about how to comply with the new testing program of Science and Social Studies content, since “it gets pushed aside from Kindergarten up, so when it comes to fifth grade and they don’t know the water cycle, it’s kinda tough” (Lucas, personal communication, April 30, 2014). There was certainly some apprehension about how to help students learn the standards-based content in these subjects now that they had missed an extensive amount of prerequisite content for the sake of Reading, Writing, and Mathematics instruction during their earlier elementary years.

Lucas referenced ideas that represent the activities of ‘using’ and ‘applying’ when he discussed using resources from the Everyday Math curriculum. He explained how he used the Essential Learning Goals from the Common Core standards as well to guide his assessment for his students. Lucas was happy with these tools and applied them to his everyday practice, noting that “the math really has it put together, like this I think is a huge tool for me. Yeah and it’s great, and my students will be able to do all this by the end of the unit” (Lucas, personal communication, April 30, 2014). Lucas prefers having resources to guide his practice, and he likes being able to incorporate tools, such as curriculum guides, to inform his instruction. This may be due to the fact that he is still gaining experience in the profession and finds external resources very helpful.

Lucas also discussed ideas that represented the activities of ‘knowing’ and ‘understanding’ when he shared his feelings regarding his level of preparedness for assessing emerging bilingual students. He conversed about how he tries to emphasize vocabulary for his students because he knows that it is helpful for emerging bilingual students. He explained, “I really push vocabulary because that’s what I was taught by the principal, like that’s what we
have to do, don’t accept, like the word ‘times’” (Lucas, personal communication, April 30, 2014). It is evident from this statement that Lucas focuses on word-level strategies to help his emerging bilingual students. This is significant because it shows that Lucas is likely unaware of strategies beyond the word level, such as vocabulary instruction, to help his emerging bilingual students in their language development. Perhaps CLD educational experience could help him to benefit his language learners in the future.

With regard to System II, Lucas mostly discussed the activity of ‘interpreting,’ as it represented some of the System II activities. He referenced ideas that represented ‘reflecting,’ since this activity represented a few of the System II activities. With equal frequency (17%), he referenced ideas that represent the System II activity of ‘concluding’ as well. He did not reference ‘comparing,’ ‘contrasting,’ ‘analyzing,’ ‘evaluating,’ ‘judging,’ or ‘integrating’ as System II activities as frequently as the aforementioned activities, if at all. This means that Lucas engages in System II activities mainly through interpreting what he thinks constitutes assessment and how he interprets assessment in his classroom. It was interesting to see that Lucas interpreted an assessment to involve “paper, pencil, [and students] sit down and take time” (Lucas, personal communication, April 22, 2014). He didn’t consider his small group interactions to be informal assessments, though he explicitly discussed how he could assess their learning from those interactions: “I can see where the kids are at with a small group, so I don’t consider those assessments, I consider that knowing their learning, which essentially that’s what an assessment is, but [laughs] I don’t know” (Lucas, personal communication, April 22, 2014). He understood the tension in his statements, which caused him to laugh. He was aware of the fact that he thinks he could better assess his students informally because “it’s easier to tell when they’re right here in front of me,” however, he had been told that this was not ‘assessment’ so he accepted and
complied with this label (Lucas, personal communication, April 22, 2014). He did not think critically to consider that assessment could also include informal interactions, because that is counter to what he was told about what constitutes assessment. This is not an uncommon reaction to assessment by current teachers in the field, since many teachers have been told that their informal assessments do not “count” as assessment.

Lucas used his System II thinking processes when he reflected on how students interpreted the content covered in class. He commented that his student asked him, “‘Are we ever going to be done with these tests?’ And then he asks, ‘Do you teach anymore?’” (Lucas, personal communication, April 30, 2014). Lucas used this as a proxy example to demonstrate how he felt about the situation. He discussed how “we haven’t been able to really teach and get together because we’re worried about reviewing for TCAP and reviewing for CMAS and reviewing for interims and so we haven’t really had the time to sit down and like ‘Hey what are we doing in this unit? What can we help each other with for this unit, in order to get to the actual assessment?’” (Lucas, personal communication, April 30, 2014). Lucas reflected on how he felt that the level of collaboration and planning for lessons suffered due to the focus on testing in his school. Lucas also incorporated System II activities when he concluded, on a similar note, that “it’s like a lot of testing that takes away from the teaching. I think it’s too much for them” (Lucas, personal communication, April 30, 2014). He came to the conclusion that there is an excessive amount of testing after he shared his reflections on the amount of testing and its effect on the students and teachers.

With respect to the CE component, Lucas mostly discussed the activity of ‘collaborating,’ as it represented some of the CE component activities. With equal frequency (23%), he referred to activities that represent ‘adapting’ and ‘transforming’ simultaneously. He also referenced
ideas that represent the activity of ‘criticizing,’ with it appearing in a few of the CE component activities. He did not reference ‘proposing,’ ‘imagining,’ ‘questioning,’ ‘organizing,’ or ‘creating’ as CE component activities as frequently as the aforementioned activities, if at all. This means that Lucas engages in CE component activities mainly through collaborating sometimes with colleagues, adapting and transforming lessons, and criticizing the lack of more extensive collaboration in his school.

Lucas showed evidence of CE component-type activities when he spoke of collaborating with colleagues. He expressed more interest in collaborating with colleagues than the other participants in the current study. This may be due to the fact that he has less teaching experience than most other teachers in the current study. He also has less support in his school to collaborate than Adele did, the first-year teacher in the study. Lucas explained that “I would love for more [collaboration], but then again I teach Mathematics, Science, and Social Studies, and the other teacher that goes, she does all the subjects, so… she’s got a lot more going on” (Lucas, personal communication, April 30, 2014). Lucas felt that he could not collaborate with the other fifth grade teacher because of the fact that she taught all subjects and also because he lamented, “she’s been here for eight years and kind of does her own thing. So it’s a bummer for me, but oh well” (Lucas, personal communication, April 30, 2014). He imagined scenarios where they could improve their collaboration on assessment specifically because he admitted that his school was just beginning to attempt alignment of assessments. He noted, “I think the more we align ourselves, because we haven’t aligned ourselves a whole lot, it’s gonna make our assessments align a lot more too” (Lucas, personal communication, April 30, 2014). He explained that they are going to try to use the same vocabulary across grade levels and that they would try to align more together vertically across grades and horizontally across grade teams.
Lucas criticized his school, which provided further evidence of CE component-type activities. He criticized the lack of collaboration with his colleagues due to testing constraints placed upon the school: “I know a big thing at this school is that there’s so much testing and we can’t, we don’t have the time, to really work with one another to really see what each other needs” (Lucas, personal communication, April 30, 2014). He mentioned that it is a collective goal for the entire school to work together more as a united group. Lucas also criticized the fact that it is a difficult school in which to teach because of the lack of available support for the students. He explained that he plays many different roles in order to support students, since “you’re not just the teacher, you’re the nurse, you get them band aids, you’re their psychologist, like ‘What happened at home?’ And then you have to deal with that kid because maybe grandma died, or maybe mom was angry at them, and so they’ll be shut down for the day unless you get them reeled back in” (Lucas, personal communication, April 30, 2014). He thinks that teachers would not usually fill these roles if there were appropriate support staff at school. Lucas likened this situation to the children’s book, “Caps for Sale,” since he must wear so many different hats when he is teaching at this school.

Lucas also showed evidence of CE component-type activities when he talked about adapting and transforming his lessons and assessments. He described how he transforms his lessons to fit the needs of his students. After introducing new content, he forms small groups with his students to help reinforce the content of each lesson. He shared how, “with my lower groups I’ll do an easier problem, so if it’s just the area I’ll do like a 3x5. But if I have my high group I will split this, the area, the rectangle into like three different parts and have them find each part to add them together so it will be more difficult for them” (Lucas, personal communication, April 22, 2014). Lucas focuses on differentiation in his classroom, then, each
time he introduces new content. He adapts assessments to the needs of his students as well. He scaffolds the assessments a bit by “allowing some students to use a calculator because they don’t have their multiplication down,” and he gives some of his students manipulatives to use during an assessment (Lucas, personal communication, April 22, 2014). In these ways, he addresses and adapts to students’ needs during assessment situations.

To summarize, Lucas engages in critical thought and action less frequently than other teachers. He engages mostly in System I type activities, such as complying with regulations, using and applying strategies he learned, and knowing straightforward ways to support his emerging bilinguals. The CE component activities in which he engaged mainly comprised of collaborating with other teachers to learn how to improve his practice, criticizing the way in which his colleagues operate, or by adapting his lessons and assessments to students’ needs. This is not necessarily surprising, since he is a relatively new teacher to the profession and is still learning the profession to a certain degree. Therefore, one would categorize his self-identified behavior as residing lower on the spectrum of critical thought and action with regard to assessment for emerging bilingual students and beyond.

*Observation*

With respect to System I specifically, Lucas mostly exhibited activity that represented ‘listening,’ as it comprised approximately half of all System I activities. He also displayed activity that represented ‘becoming familiar’, with it appearing in some of the System I activities. He did not exhibit ‘knowing,’ ‘understanding,’ ‘using,’ ‘applying,’ ‘complying,’ ‘assuming,’ or ‘copying’ as System I activities as frequently as the aforementioned activities, if at all. This means that, from the classes observed, Lucas engages in System I activities through listening to his students and becoming familiar with their learning needs.
He listened to his students when, for example, he asked the class to turn and talk to a partner about what they learned in a video on hurricanes and tornadoes. It was evident that he was listening to the content of students’ conversations, as he redirected students’ discussions when they were off-task (Lucas, field notes, April 30, 2014). Lucas became familiar with his students’ learning needs when he asked students to complete an ‘entrance ticket’. The entrance ticket listed several equations to solve, all of which were relevant to the current unit of study. He explained to students that he wanted to “see what you know” (Lucas, field notes, May 6, 2014). He became familiar with what students knew about the current unit also when he asked students to give him a ‘thumbs up or thumbs down’ with regard to their understanding of the content covered that day (Lucas, field notes, May 6, 2014).

With respect to System II specifically, Lucas only exhibited activity that represented ‘judging,’ as it comprised the only System II activity observed. He did not exhibit ‘comparing,’ ‘contrasting,’ analyzing,’ ‘interpreting,’ ‘evaluating,’ integrating,’ ‘reflecting,’ or ‘concluding’ as System II activities at all. This means that, from the classes observed, Lucas engages in System II activities through judging his students’ responses in class. Lucas judged his students responses in the context of a game the class was playing as a way to review for the interim test. For each review problem, students would respond individually at first, then with their small group, and finally as a class. The teacher allotted formal points to each student group based on their explanation of how they solved the problem, the math vocabulary they used in their explanation, and whether or not their answer was correct. The formal points in the game were based on Lucas’ professional judgment.

With respect to the CE component specifically, Lucas mostly exhibited activity that represented ‘adapting’ and ‘transforming,’ as this pair of activities comprised nearly all of the
CE component activities. He also displayed activity that represented ‘collaborating,’ with it appearing in a small percentage of the CE component activities. He did not exhibit ‘proposing,’ ‘imagining,’ ‘organizing,’ ‘questioning,’ ‘criticizing,’ or ‘creating’ as CE component activities at all. This means that, from the classes observed, Lucas engages in CE component activities through adapting and transforming his instruction to meet students’ needs and through collaborating with students in the assessment process.

Lucas adapted and transformed his instruction often when he worked in small groups with students. He listened to students’ thinking about a problem and responded by providing feedback to improve their problem solving skills. For example, Lucas asked a student: “Why do you say seven [marbles]?” (Lucas, field notes, May 5, 2014). He also provided encouragement when he noticed an emerging bilingual student struggling with the material: “I want to recognize that this is getting difficult and you can’t give up. I need some more effort out of you” (Lucas, field notes, May 6, 2014). This instruction was tailored to the individual student with whom he was working. Lucas also collaborated with students in the assessment process by engaging his class in peer assessment. Specifically, when students played a review game to study for the interim test, Lucas asked the class to give a ‘thumbs up or thumbs down’ as a reaction to the answer that was volunteered. If a student had a ‘thumbs down,’ they needed to explain why they disagreed with the other student’s answer and how they would recommend solving the problem instead (Lucas, field notes, April 30, 2014). Lucas also awarded formal points to students for their answers, but he asked the class to provide their opinions on the answers first through the ‘thumbs up or thumbs down’ technique. In this way, students engaged in peer assessment and became active participants in the assessment process.
To summarize, Lucas engages in critical thought and action rather infrequently in observed classroom situations. Therefore, one would categorize him as residing on the lower end on the spectrum of critical thought and action with regard to observed activity in the area of assessment for emerging bilingual students.

Serena

Interview

With respect to System I specifically, Serena mostly discussed the activity of ‘complying’, as it represented exactly half of all System I related activities. She referenced the activities of ‘knowing’ and ‘understanding,’ which simultaneously appeared in a few of the System I activities. Also appearing in a few of the System I activities were both ‘listening’ and ‘becoming familiar.’ She did not reference ‘using’, ‘applying,’ ‘copying,’ or ‘assuming’ as System I activities as frequently as the aforementioned activities, if at all. This means that Serena engages in System I activities mainly through complying with regulations.

One main reason that compliance was a theme for Serena was due to grading procedures. As in Valera’s school, Serena’s school grading policy also mandates that teachers give students grades that reflect only their standards-based test scores. However, in Serena’s school, the students’ grades are based on the scores from the tests modeled after the state standardized tests, which are given every six weeks in the charter school. In Valera’s school, the grade is based solely upon the annual state standardized test score. Though Serena’s school allows for additional test scores to be included in the students’ grades, there is still no opportunity for any classwork, quizzes, homework, or participation to factor into the students’ grades. Serena expressed difficulty explaining to her students and their parents how much they have grown, “because their grades are only a test, it’s really hard” (Serena, personal communication, May 16,
There is a tension between System I and CE component activities in this situation, since all the teachers are asked to write the tests that are given to students every six weeks, the scores from which become their grade. Creating a test is a CE component activity, yet being required to submit the score as the only source of a student’s grade is an act of compliance on the part of the teacher, thereby involving a System I activity as well.

The activity of ‘complying’ also plays a role in Serena’s self-identified behavior when she discussed the state standardized testing system. She expressed frustration over the fact that the type and relative frequency of various content standards appearing on the tests is not shared with teachers, since she is unable to tell what to expect on the tests. She sighed, “But I don’t have any control over that. And the score that they get is the score that they get, and the score their school gets” (Serena, personal communication, May 16, 2014). This issue that the contents of the test are both unknown to teachers and out of their control is a salient one, since it appeared as a theme in the surveyed teachers as well as in the interviewed teachers.

Serena discussed ideas that represent the activities of ‘knowing’ and ‘understanding’ simultaneously when she shared how she urges students to show what they know when she assesses them. She wants to better know and understand how students think during the process of assessment instead of simply finding an answer and saying, “I’m done” (Serena, personal communication, May 16, 2014). She invests a lot of time in collecting information to understand what students know about the content and the language of the curriculum.

Serena also shared ideas that represent the activities of ‘becoming familiar’ and ‘listening’ when she explained how she listens to her students giving her “blank stares” or telling her “I have no idea how to do the math, because I don’t know the words” (Serena, personal communication, May 16, 2014). She posited that when she sees these reactions from her
students, “they shut down very quickly” (Serena, personal communication, May 16, 2014). She has become familiar with how to confront what she termed ‘learned helplessness’ on the part of her students. She adapts her instruction to her students’ needs when this learned helplessness takes place, and she helps her students learn the challenging vocabulary in various ways so they can feel confident in learning the subsequent content objectives. Serena is only able to respond to her students’ needs because she listens so well to her students when they become confused. This is an important part of formative assessment, when teachers need to collect information about what students understand in order to tailor their instruction to their needs, whether they are language- or content-based.

With regard to System II, Serena mostly discussed the activity of ‘reflecting,’ as it represented some of the System II activities. She also heavily referenced ideas that represent the activity of ‘concluding,’ with it appearing in a few of all System II activities. With the same frequency (14%) she referenced ‘contrasting’ and ‘interpreting’ in a few of the System II activities. She did not reference ‘comparing,’ ‘analyzing,’ ‘evaluating,’ ‘judging,’ or ‘integrating’ as System II activities as frequently as the aforementioned activities, if at all. This means that Serena engages in System II activities mainly through reflecting on her practice, making conclusions about how her students learn, contrasting ways to assess her students, and interpreting how her students experience the testing procedures in her school.

One recurring theme in her reflection was with regard to project-based learning. Serena discussed how she prefers implementing project-based learning to achieve two main aims: to increase student motivation and to introduce social justice themes into the curriculum. Serena reflected on projects that she had already completed this year with her students, such as giving students a budget and planning a menu for their school trip to Utah or creating story books of
geometry vocabulary that they would then use to share with younger students. She explained that these projects were effective in heightening student motivation because “they had some choice in the matter, which gave them more buy-in first of all, but then also they got to show so much more of their knowledge because they were into it and they were creating a product that was bigger than paper / pencil, I’m by myself” (Serena, personal communication, May 21, 2014). Collaboration and authenticity of the projects played a big role in helping students to become interested in the content, in Serena’s opinion.

She also reflected on how she used these projects to increase social justice themes in the curriculum. She argued, “who cares… about percentages that are made up from my brain, I mean, really, who cares? But if it’s talking about their neighborhood versus other neighborhoods, or it’s talking about income potential or making a budget,” then that is different (Serena, personal communication, May 16, 2014). Serena thought that it is much more meaningful to students’ experience if her lessons have the potential to transform their lives as well as helping them to access the learning goals. She continued, “They can go home and talk to their parents about it, and their parents can participate in that discussion. That changes their engagement level, but also changes their reasoning for wanting to get the right answer, and their reasoning for wanting to be able to explain the right answer” (Serena, personal communication, May 16, 2014). Serena reflected on instructional practices that are projected to increase students’ motivation and allow for a transformative educational experience. This extends to a CE component-type activity when she proposes these social justice-oriented objectives for her students.

Serena discussed ideas that represent the activity of ‘concluding’ when she posited that her sixth graders are at a point where “the abstract math thinking is really hard” (Serena,
personal communication, May 16, 2014). She made this conclusion based on her experience in the classroom for the past few years, working with sixth, seventh, and eighth graders. She noted that few sixth grade students have arrived at a place developmentally where they can think in an abstract manner. Based on this conclusion, she decided that she needs to help “supplement by as many visual processes, rote memorizations as I can” in order to help her students tackle the abstract concepts present in the sixth grade curriculum (Serena, personal communication, May 16, 2014).

Serena also shared ideas that represent the activity of ‘contrasting’ when she discussed how she likes to assess her students and what kind of knowledge she wants them to have. She stated that she would like for students to focus on attaining the kind of knowledge that would allow them to have a full understanding of the Mathematics concepts instead of being test savvy. She said, “I’d rather them know how to do the math than to actually get the right answer; like, how to do the math. Understanding what they’re doing is so much more important to me” (Serena, personal communication, May 16, 2014). Serena prioritizes students’ deep content understanding, as opposed to being able to select a correct answer from a list of possible answers.

She also contrasted the way in which she likes to assess her students. She argued that “group assessments, it’s harder to see [what each student knows]. A lot more of it’s verbal” (Serena, personal communication, May 16, 2014). However, she explained that she likes to assess students as they work in small groups because “all of my class just does so much better as a group of learning and understanding, especially as so many of them are intermediate to advanced emerging bilinguals, like the discussion piece, and the helping process of orally, both in English and Spanish, being able to just have the discourse, not just written on paper that’s so
isolated helps them figure out what’s going on” (Serena, personal communication, May 16, 2014). She contrasted the relative advantages and disadvantages of assessing students as they work in groups.

Serena discussed ideas that represent the activity of ‘interpreting’ when she shared her interpretation of how emerging bilingual students experience the testing that they undergo in school. She argued that learning third through sixth grade fractions in six weeks and then being tested on it would be overwhelming for anyone. She continued, “especially just the fatigue of taking a 45 question test with all sorts of different language, and having to remember all of these different things… it’s really hard!” (Serena, personal communication, May 16, 2014). She used this interpretation to help her implement CE component-related activities to adapt her assessment process, such as helping students to track their progress so they can feel more successful despite some challenges posed by the structure of testing in the school.

Additionally, Serena referenced the activity of ‘interpreting’ when she explained her view on the state standardized testing program and whether her students’ Mathematics scores reflect their progress in her Mathematics class. She gave her interpretation, which was that her students “can do all the math, but without that, like, nudge of like, ‘Remember, you know this!’ They just, they’re not able to show everything that they know” (Serena, personal communication, June 4, 2014). She continued to detail how she emphasizes application of concepts in her class, because she thinks her students will be able to show their understanding better. Serena went on:

And if they have multiple experiences with those kinds of things, then they’ll be able to show their proficiency better on TCAP but also on my tests and also just use their math moving forward, which is really what I care about. (personal communication, June 4, 2014)
Her interpretation of how important it was for her students to be able to apply mathematical concepts is interesting because it shows that she cares most about students’ ability to use math going forward in their lives for authentic purposes. She does not simply consider the external pressures on her, directing her to focus on application for the purpose of improving her students’ standardized test scores. Instead, she claims to focus on application for the sake of helping students to understand math better in any context. This lends itself to a CE component-type activity as well, because she teaches in response to her critical thinking about what students will need as they move forward in their lives instead of teaching in a certain way because it is required.

To segue into a full analysis of her CE component activities, Serena mostly discussed the activity of ‘creating,’ as it represented some of the CE component activities. She also referenced ideas that represent the activities of both ‘adapting,’ and ‘transforming,’ as they appeared in some of the CE component activities. She also referenced ideas that represent the activity of ‘criticizing,’ with it appearing in a few of the CE component activities. She did not reference ‘proposing,’ ‘imagining,’ ‘questioning,’ ‘organizing,’ or ‘collaborating’ as CE component activities as frequently as the aforementioned activities, if at all. This means that Serena engages in CE component activities mainly through creating assessment instruments and alternatives, adapting and transforming her assessment to fit students’ needs, and criticizing the school culture on their lack of attention paid to issues of emerging bilingual students.

Specifically, Serena created assessment alternatives as a response to the fact that the tests every six weeks were the only products that contributed to the final grade for her students. Serena covertly found ways to add points to students’ grades for ‘assessment items’ by including projects they worked on during the course of the semester. She explained that she gave points for
such an assessment item “because it was an assessment. And it was a demonstration of their knowledge. It just wasn’t pencil and paper at that very day and time. And so, no one’s talked to me about it, yet” (Serena, personal communication, May 16, 2014). By saying this, Serena was insinuating that an administrator could come and confront her about this behavior and ask her to refrain from it. Serena bent the rules because she felt it was best for her students, which is a CE component-type activity. She did not simply follow orders; instead, she thoughtfully considered the consequences for her students if she only included the larger assessments and decided it was better for their development to sneak in a few project assessment points to their overall grade. This is the kind of behavior one would expect from someone who takes action in response to their critical thinking about an issue. This is an even more admirable action when one considers its potential negative repercussions.

Serena created assessment alternatives for her students as a response to the requirement to have only the test scores factor into students’ grades. Though it does not count towards their grade, Serena sends home quizzes, exit tickets, and positive notes of what happened in class. She helps students and parents track their progress even if it is not reflected in their test score, and therefore not reflected their grade: “So, helping them track their progress, too, of ‘Oh, I’m growing!’ And then when I’m having parent conversations, as well, this is where I saw them grow” (Serena, personal communication, May 16, 2014).

Serena discussed how encouraging students to track their progress also helps them to maintain or build a positive self-concept with regard to the academic realm. Serena argues that: “[It’s [the test given every six weeks] much more of what they know, and what they’re able to demonstrate at that point in time, which makes it a lot higher stakes. And it contributes to our helplessness of, ‘I always am unsatisfactory, I must not know anything.}
I’m always failing.’ So that’s why I’m trying to build in as much as I can of the tracking of ‘Where have you grown?’ ‘What have you done better at?’ ‘How do you learn, what have you learned this year, what have you learned in these six weeks?’ Just to try and temper that. But it’s tough…” (Serena, personal communication, May 16, 2014).

Positive self-concept is extremely important to students’ academic success, particularly as students are developing their identity in these early years of schooling (Dweck, 1999). For emerging bilinguals, this is even more critical because learning a language is much easier if one’s affective filter is low in the learning environment (Richard-Amato, 2010). Alma communicated similar ideas and also developed ways to help students recognize their growth with regard to the learning goals. Serena discussed how she wanted students to engage in project-based learning so they could take ownership and pride over their learning, thereby building positive self-concept. She expressed that she created projects and managed to quietly incorporate them into her assessment system. Her intention was that students would think to themselves: ‘I want to create a better product. People are going to see it. I’m proud of it,’ versus their test fatigue that I feel like a lot of our kids are under right now where they’re just like, ‘Oh my gosh, another test. I don’t want to take another test,” (Serena, personal communication, May 21, 2014).

Serena represented the CE component activities of ‘adapting’ and ‘transforming’ when she talked about how she shapes her instruction to fit the linguistic needs of her students by engaging in project-based learning. She explained how it was optimal for her to attach content covered in the curriculum to project-based learning because students would be encouraged to discuss the math much more in depth. Thus, they practiced the language as well as the content of Mathematics through the authentic context of the projects. She discussed how “the experience of talking about it a whole bunch of times gives them a confidence and gives them an ability to
express themselves in a way that they don’t have with just paper/pencil experiences (Serena, personal communication, May 21, 2014). It is clear from this statement that Serena also liked projects because it decreased the affective filter and increased positive self-concept for language learners. She mentioned how she also uses manipulatives, sentence stems, explicit vocabulary practice, and word problem practice all in order to specifically help language learners access the content. Therefore, Serena made many adaptations to help increase her emerging bilingual students’ access to math language in addition to math content. She also noted how she adapts her questioning styles to specific students who might need more scaffolding in order to give them more support (Serena, personal communication, May 21, 2014). Serena employed many techniques to adapt and transform her instruction to fit the needs of her students.

Serena engaged in criticism as a CE component-type activity as well. She criticized her colleagues for not focusing on how to provide emerging bilingual students access to the kind of language found in the assessments administered in her school. She noted that they pull items for their tests from assessment banks which are not designed for language learners. She critiqued: “I feel like we’re not engaging in a sort of dialogue about how are our students really accessing this test and what are they doing?” (Serena, personal communication, June 4, 2014). Serena expressed interest not only in providing her students access to language in assessment, but also creating a community with her colleagues where they could share ideas about how to accomplish this goal.

To summarize, Serena engages in critical thought and action frequently mainly by creating assessment alternatives, adapting and transforming her assessments to her students’ linguistic needs, and criticizing colleagues for their lack of attention to language when addressing assessment. She engaged in some of these activities despite the fact that these
activities were not encouraged by the school administration, which speaks to her drive and commitment to engaging in these activities. Therefore, one would categorize her self-identified behavior as residing high on the spectrum of critical thought and action with regard to assessment for emerging bilingual students and beyond.

**Observation**

With respect to System I specifically, Serena mostly exhibited activity that represented ‘becoming familiar,’ as it comprised approximately half of all System I activities. She also displayed activity that represented ‘listening,’ with it appearing in some of the System I activities. She did not exhibit ‘knowing,’ ‘understanding,’ ‘using,’ ‘applying,’ ‘complying,’ ‘assuming,’ or ‘copying,’ as System I activities as frequently as the aforementioned activities, if at all. This means that, from the classes observed, Serena engages in System I activities mainly through becoming familiar with what students understand with respect to the content objectives and by listening to students construct knowledge during class.

She became familiar with students’ understanding of the content objectives by taking informal surveys of students’ progress on the classroom tasks. She also asked students their opinion on the answers to the math problems. For example, she asked students “Alright, show of hands – how many think Plan B is better? Who thinks that Plan A is better?” She was monitoring that everyone participated in this survey because she noted that four students still needed to raise their hands in response to her question. She listened to students construct knowledge during class by circulating among every pair of students as they engaged in group work. It is clear that she listened to students as she circulated around the room, since she reported what she heard in students’ conversations to the whole group. She noted “I’m having trouble [because I’m] finding that people are talking without solving it” (Serena, field notes, May 29, 2014). She continued
providing feedback based on what she specifically heard when she listened to students discuss in their groups.

With respect to System II specifically, Serena mostly exhibited activity that represented ‘integrating,’ as it comprised approximately half of all System II activities. She displayed activity that represented ‘comparing,’ with it appearing in a small percentage of all System II activities. She also displayed activity that represented ‘judging,’ as it was found in a small percentage of all System II activities. She did not exhibit ‘contrasting,’ analyzing,’ ‘interpreting,’ ‘evaluating,’ ‘reflecting,’ or ‘concluding’ as System II activities at all. This means that, from the classes observed, Serena engages in System II activities mainly through integrating real life experiences into her lessons, comparing vocabulary for students to develop linguistically, and judging students’ responses in class.

She integrated real life experiences into her lessons when she orchestrated a class activity in which students needed to evaluate various cell phone plans based on cost and characteristics of the plans. She made linguistic comparisons to help students better understand content material. For example, she asked students to tell her the vocabulary word to indicate the act of taking a variable out of an equation and putting a number in its place. They struggled to find the target vocabulary word, so the teacher waited and then hinted, “Kind of like when you take one teacher out and put another one in” (Serena, field notes, May 21, 2014). With this information, the students were able to compare a substitute teacher to the mathematical substitution they were asked to do in the problem. This comparison on the part of the teacher helped students to make linguistic connections themselves. Serena also judged students’ responses as correct or incorrect with explicit corrective feedback. She did this by verifying orally when she viewed their responses on individual whiteboards. She walked around her classroom, remarking “Yes,”
“Perfect,” and “Yup.” The teacher’s quick and explicit judgment as a form of feedback may help students to track their own learning progress.

With respect to the CE component specifically, Serena mostly exhibited activity that represented ‘adapting’ and ‘transforming,’ as this pair of activities comprised most of the CE component activities. She also displayed activity that represented ‘creating,’ with it appearing in some of the CE component activities. She did not exhibit ‘proposing,’ ‘imagining,’ ‘collaborating,’ ‘organizing,’ ‘questioning,’ or ‘criticizing’ as CE component activities as frequently as the aforementioned activities, if at all. This means that, from the classes observed, Serena engages in CE component activities mainly through adapting and transforming her classroom instruction based on her students’ reactions and needs. She also engages in CE component activities through creating a classroom environment that promotes meaningful, linguistically challenging classroom discourse and meta-cognitive connections.

Serena adapted and transformed her instruction based on her students’ needs when she modified activities after noticing her students struggling with a task. For instance, when she noticed students struggling with an activity in which they were to evaluate various cell phone plans, she exclaimed, “I am hearing from you that you’d like some more support. So let’s put the expression in the first part, and then let’s explain who the plan would be good for and who it would not be good for” (Serena, field notes, May 22, 2014). She continued to provide more structure for the activity by suggesting a starting point for students and by adding examples until her students could confidently engage in the activity. She even modified the kind of product that she expected students to produce in one class period. At the end of the cell phone plan activity, Serena noticed that students were still not moving as smoothly through the activity as she had anticipated. She then told her class, “I’m going to adjust. Here’s what I’d like you to produce
today. I want you to turn into me a binder page…. We didn’t finish doing the math for all of the bills” (Serena, field notes, May 22, 2014). Serena excelled at being forthright with her class by explicitly telling students as a group what she learned about their needs during group work. More than once, she explained that she had a specific plan for the day, but was going to adjust it based on what students needed to be successful in the day’s activities.

Additionally, she created an environment where students could interact by engaging in meaningful, linguistically challenging discourse when she emphasized students’ explanation of the math problems instead of simply focusing on the calculations. In a continuation of the cell phone plan activity, she told her students, “I’d like to hear two different word answers. Explanations of why neither price was cheaper or they both have the same price” (Serena, field notes, May 29, 2014). She also asked students to develop an argument to defend their position, which required students to engage in discourse for an authentic purpose of evaluating and selecting a superior cell phone plan. She questioned the class, “OK, who would like to give their explanation of why they think that Plan B is better?” Serena, field notes, May 29, 2014). Students needed to use persuasive language to convince others that one cell phone plan was a better choice than the others. Constructing such an argument is linguistically as well as mathematically challenging.

She also created an environment where students could easily create meta-cognitive connections for themselves. For example, Serena asked students to work on creating a “resource book,” which is a book written by students that contains strategies for problem solving. The purpose of the resource book is to facilitate students’ studying and to help students to recall the problem solving processes in the future. This aids students in understanding how to monitor their own learning. When Serena created ways for students to make these meta-cognitive connections,
she provided them with a powerful tool that students will be able to employ throughout their educational experience.

To summarize, Serena engages in critical thought and action somewhat frequently by adapting and transforming her lessons to her students’ needs and by creating challenging opportunities for discourse and meta-cognitive connections in observed classroom situations. Therefore, one would categorize her as residing higher on the spectrum of critical thought and action with regard to observed activity in the area of assessment for emerging bilingual students.

Claire

*Interview*

With respect to System I specifically, Claire mostly discussed the activity of ‘complying,’ as it represented approximately half of all System I related activities. She also heavily referenced ideas that represent the activities of ‘knowing’ and ‘understanding,’ with them appearing simultaneously in some of System I activities. She did not reference ‘becoming familiar,’ ‘listening,’ ‘using,’ ‘applying,’ ‘copying,’ and ‘assuming’ as System I activities as frequently as the aforementioned activities, if at all. This means that Claire engages in System I activities mainly through complying with school regulations and by having knowledge and an understanding of issues pertaining to language and culture in assessment.

She discussed compliance with school regulations with regard to assessment of emerging bilinguals when she explained her school-wide assessment policies. The school in which Claire teaches administers tests in each subject every six weeks. These tests represent a cumulative assessment of students’ progress in each subject over the course of the unit covered each six week period. Then, there is also a pre- and post-test developed by the charter network that is given to students in subjects that are not tested by the state standardized testing program. Since
sixth grade Science is not a subject that is tested by the state assessment program, Claire’s students are tested once at the beginning of the year with the pre-test and once at the end of the year with the post-test. Claire shared that these tests are the only tests that are completely external to her control. She explained, “I don’t have knowledge when we give the Science pre- and post-test,” and in fact she is not given access to create, modify, or even know the content of the items on that test (Claire, personal communication, May 5, 2014).

Claire also described how she must come in during the summer break to collaborate with colleagues to create the tests that will be given to students every six weeks during the school year. Although this event involves CE component activities, they are forced activities. Thus, it is somewhat difficult to piece apart the level of agency and initiative that Claire would have if these CE component-type activities were not mandated by the school. She also explained how, though she works with her colleagues to create the tests given to students every six weeks, “you have until three weeks before the test and then it’s locked, and we can’t make any edits to the test during those three weeks before” (Claire, personal communication, May 5, 2014). This regulation is to help increase the validity of the test, Claire explained, in case someone realized they failed to cover some targeted content on the test and then attempted to remove that content from the test. Although the reasoning behind this regulation reveals a process that has been critically examined, this does not reflect critical thought and action on the part of Claire as an individual, since she has no control over the process.

Claire had ideas that represent the activity of ‘knowing’ and ‘understanding’ when she shared her understanding of bias in testing with regard to emerging bilingual students. She explained how “there’s been quite a few studies on how white kids tend to do better on standardized tests and a lot of times there are cultural biases within the tests” (Claire, personal
communication, May 5, 2014). She gave the example of a test item that references an aquarium and declared that this type of item carries the assumed experience of an aquarium as well as knowledge of animals that can be found within the aquarium. Claire’s understanding of bias in testing demonstrates her knowledge of how cultural background can affect students differently in testing situations.

With regard to System II, Claire mostly discussed the activity of ‘reflecting,’ as it represented approximately half of all System II activities. She referenced ideas that represent the activity of ‘analyzing,’ with it appearing in a few of System II activities. She also referenced ideas that represent the activity of ‘evaluating,’ as it appeared in a few of System II activities. She did not reference ‘comparing,’ ‘contrasting,’ ‘interpreting,’ ‘judging,’ ‘integrating,’ or ‘concluding’ as System II activities as frequently as the aforementioned activities, if at all. This means that Claire engages in System II activities through reflecting on her practice, analyzing student performance on content and process objectives, and evaluating students’ Science vocabulary, for example.

Claire expressed ideas that represent the activity of ‘reflecting’ when she shared her goals for growth for that school year. For example, she mentioned that metacognition is “an area that I’ve been working on this year,” where she encourages students to assess their assessments and reflect on how they performed (Claire, personal communication, May 5, 2014). Claire reflected on her goal of increasing her students’ reflections on assessment. Asking students to increase their metacognitive processes in learning was a goal that Claire’s instructional coach encouraged. Claire took agency in this goal once it was suggested, as she explained how the instructional coach inspired her to “allow kids to look into their own learning rather than just me telling them, this is how you did. So that’s been one of my like areas of growth to push myself on this year.”
Claire is cognizant of improving her practice and she incorporated suggestions from her instructional coach to push her practice to a more advanced level.

Claire also expressed ideas that represent the activity of ‘analyzing’ when she discussed working with student data extensively when she participates in a ‘data day’ following each unit test every six weeks. On this day, students do not come to school. Instead, teachers spend an hour and a half of the day in professional development, then they use the rest of the day analyzing student performance and using that data to adjust their instruction. She recounted how “we spend about an hour and a half doing data analysis,” using spreadsheets that automatically calculate trends of student performance on the tests (Claire, personal communication, May 5, 2014).

Claire also shared how she analyzes student performance with regard to process objectives in addition to content objectives. She explained how process objectives include activities that are linked to the content area goals, which complement the mastering of content area goals. Claire then tracks and analyzes process objectives such as “how many times kids didn’t respond in a complete sentence, or count how many, how long it took us to get into groups, or how much work got done when we were in groups” (Claire, personal communication, May 5, 2014). Claire therefore analyzes elements of language as well as content when she tracks process objectives, since she considers issues like whether students are using complete sentences.

Claire continued to focus on language when she expressed ideas that represent the activity of ‘evaluating.’ She explained how she “will give a vocabulary specific quiz once a period typically to assess their knowledge of those words” (Claire, personal communication, May 1, 2014). This comment shows that she evaluates her students’ language skills at the word
level during each unit. Though there is more to language learning than individual words in isolation, Claire’s initiative in evaluating her students on vocabulary shows an awareness of language as an important element of Science content. She understands that her students need to master both the language and content components of Science, and focuses on evaluating both aspects in her practice.

With respect to the CE component, Claire mostly discussed the activities of ‘adapting’ and ‘transforming’ simultaneously, as they represented some of the CE component activities. With equal frequency (31%), she also referenced ideas that represent the activity of ‘creating.’ She referenced ideas that represent the activity of ‘collaborating’ as well, as it appeared in some of the CE component activities. She did not reference ‘proposing,’ ‘imagining,’ ‘criticizing,’ ‘questioning,’ or ‘organizing’ as CE component activities as frequently as the aforementioned activities, if at all. This means that Claire engages in CE component activities mainly through adapting and transforming her lessons according to students’ needs, creating assessment tools to help meet emerging bilingual students’ language demands, and collaborating with her colleagues.

Claire expressed ideas that represent the activities of ‘adapting’ and ‘transforming’ when she recounted how, on her school’s ‘data day’ every six weeks, she and her colleagues will analyze how her students performed on the test, and then they work together as a content team “making adjustments for next year, and then making adjustments for the upcoming unit.” (Claire, personal communication, May 5, 2014). She also explained how “we have about an hour and a half of individual work time to figure out how you’re going to teach students the next day the stuff that they missed” (Claire, personal communication, May 5, 2014). Claire mentioned that it was mandatory to reteach content objectives to their class if a certain number of students
incorrectly answered the item that targeted that content on the test. There is an odd dynamic of mandated practice of critical activities. It is good practice to adapt one’s instruction to students’ needs, but if teachers are mandated to do so, it is difficult to determine how much of this stems from a teacher’s individual initiative.

Claire communicated ideas that represent the activity of ‘creating’ when she discussed the packets of worksheets for classwork that she gives to students each week. She explained that the packets have the content material that cover the objectives targeted each week. She described how she creates accommodated and modified packets to help scaffold and improve student learning. She explained that the accommodated packets target emerging bilingual students by addressing some of the language demands of the lessons, whereas the modified packets are created to adapt lessons for students with special needs. She shared that the accommodated packets have “filled-in information for them already or little notes, or it’ll have more pictures for vocabulary words” (Claire, personal communication, May 1, 2014). This shows that Claire takes time to create supports for her emerging bilinguals by tailoring a different version of the classwork packet to their language needs. Since classwork and homework comprise 50% of students’ grades in Claire’s school, the accommodated classwork packets serve as assessments that are accommodating to linguistic needs.

Claire also expressed ideas that represent the activity of ‘collaborating’ when she recounted how she works together with her colleagues to plan for instruction and assessment with emerging bilingual students. On the school’s ‘data day,’ teachers analyze data of student performance on the unit test that was given the day before and they plan how to adapt subsequent instruction. Teachers do this analysis and planning as a content team for the data day, which requires significant collaboration among colleagues. She also detailed her leading role as a
collaborator with her colleagues when she talked about her role as a curricular touch point for her team. She explained how, in the previous year, she would “post daily materials that people could access, and we had to teach the same objectives every day” (Claire, personal communication, May 5, 2014). Claire recounted a shift in the way she and her colleagues collaborated over the past year, and she described how this year it was a lot more individualized with the unit test serving as the only point of alignment. She explained that, for next year, “the belief is that even if you want to do your own track, there will be centralized resources to use if you’d rather just stay with kind of the group plan” (Claire, personal communication, May 5, 2014). It is interesting to hear that, even in Claire’s CE component-type activities of collaboration, there are undertones of compliance as a school. Claire and her colleagues “had” to teach the same objectives every day in the previous year, and that this year “it’s a lot more individualized,” which brings the observer to realize that the school policy and culture imposed individualism rather than the teachers themselves (Claire, personal communication, May 5, 2014). When she mentioned “that was the belief as a network last year, and then it all, it was a total kind of paradigm shift at the network level this year,” it becomes clear that Claire’s school imposes guidelines for the ways in which teachers are expected to collaborate. This makes it difficult to discern how much of Claire’s collaboration with her colleagues on assessment and subsequent adaptation to instruction can be attributed to Claire herself and how much can be attributed to her response to school policies. Claire may very well engage in these CE component-type activities on her own, but it is impossible to disentangle which activities arise from her own sense of agency and which activities arise from her compliance with school policy.

To summarize, Claire engages in CE component types of activities slightly less frequently than other teachers in the current study. Therefore, one would categorize her self-
identified behavior as residing in the middle of the spectrum of critical thought and action with regard to assessment for emerging bilingual students and beyond. However, she would score quite high on System II-type activities, since she focused her energy heavily on analysis of student performance. This was partly due to the school’s focus on data analysis, as the school asked teachers to prioritize analysis of student data. As a whole, it was difficult to discern which of Claire’s activities were carried out of her own volition and which activities were carried out due to her compliance with external forces, such as her school’s policies.

Observation

With respect to System I specifically, Claire mostly exhibited activity that represented ‘listening,’ as it comprised approximately half of all System I activities. She displayed activity that represented ‘becoming familiar,’ with it appearing in some of the System I activities. She also displayed activity that represented ‘knowing,’ as it was visible in a small percentage of all System I activities. She did not exhibit ‘understanding,’ ‘using,’ ‘applying,’ ‘complying,’ ‘assuming,’ or ‘copying,’ as System I activities as frequently as the aforementioned activities, if at all. This means that, from the classes observed, Claire engages in System I activities mainly through listening to her students, becoming familiar with their needs, and knowing the learning objective for the day.

For instance, Claire listened to her students when she asked students which parts of a quiz they struggled to answer, and then she focused on students’ responses. She listened to what students found challenging about the quiz both right after they took the quiz and when they reviewed the quiz as a class (Claire, field notes, May 7 and 12, 2014). She also became familiar with the areas in which students struggled when she asked students to raise their hands if they struggled with the bottom section of a warm-up quiz called a “Do Now” (Claire, field notes, May
Claire also knew what the learning objective was for each observed class period when she posted it on the whiteboard. She made the learning objective evident for her class as well by asking them to read it aloud so they would be aware of the learning objective each class period (Claire, field notes, May 7, 12, and 13, 2014).

With respect to System II specifically, Claire mostly exhibited activity that represented ‘integrating,’ as it comprised most of the System II activities. She also displayed activity that represented ‘judging,’ with it appearing in some of the System II activities. She did not exhibit ‘comparing,’ ‘contrasting,’ analyzing,’ ‘interpreting,’ ‘evaluating,’ ‘reflecting,’ or ‘concluding’ as System II activities at all. This means that, from the classes observed, Claire engages in System II activities through integrating real life context into her instruction and through judging her students’ responses. She integrated authentic experiences into her instruction when she assigned students a class activity in which students would create a “fish book” profile for a specific animal in a food chain. The “fish book” profile was specifically modeled to resemble a Facebook profile so that students would be able to relate the class activity to their lived experience with the social media application. This integration of real life tools, such as the Facebook social media tool, helped to motivate students who enjoy using Facebook. It also helped some students focus on the content, since the context of the activity was readily familiar to them. Additionally, Claire judged students’ responses when she provided explicit corrective feedback to students about the content. She answered some students in the affirmative with a “Yeah!” (Claire, field notes, May 13, 2014). This provided students immediate feedback through a judgment of their response.

With respect to CE component specifically, Claire mostly exhibited activity that represented ‘adapting’ and ‘transforming,’ as this pair of activities comprised most of the CE
component activities. She also displayed activity that represented ‘questioning,’ with it appearing in some of the CE component activities. She did not exhibit ‘proposing,’ ‘imagining,’ ‘collaborating,’ ‘organizing,’ ‘criticizing,’ or ‘creating’ as CE component activities, if at all. This means that, from the classes observed, Claire engages in CE component activities through adapting and transforming her instruction based on students’ needs and by posing meta-cognitive questions to her students.

Claire adapted and transformed her instruction when she, for example, asked students why a group of organisms would be important to an ecosystem (Claire, field notes, May 12, 2014). When students volunteered their answers, she wrote down students’ responses on the whiteboard for the whole class to see and she used it as a springboard for further discussion. Claire reacted to what students responded and she modified her subsequent classroom conversation accordingly. She also posed meta-cognitive questions to students when she asked students which items on the quiz she found to be challenging (Claire, field notes, May 12, 2014). She could have posed a deeper inquiry to find out more about students’ thinking by asking students why they found particular items to be challenging, and she might have discovered that some of their challenges were related to language. However, she did address students’ meta-cognitive processes, which helps students to learn to be critical and monitor their own learning. This represents a more critical action on the part of the teacher, since she is initiating these meta-cognitive reflections among her students.

To summarize, Claire engages in critical thought and action rather infrequently in observed classroom situations. Therefore, one would categorize her as residing on the lower end of the spectrum of critical thought and action with regard to observed activity in the area of assessment for emerging bilingual students.