

English Language Learners and Measures of Academic Progress (MAP) Testing

Rachel E. Bohlman

University of Wisconsin - Whitewater

A Project Submitted in Partial Fulfillment
Of the Requirements of the
Master of Science-Curriculum and Instruction

This project was approved by:

Melissa Freiberg

Acknowledgements

I would like to thank several colleagues in the school district including Lisa Marek, Melody Leung, Ann Schoenberger, and the teachers I interviewed for the project. My advisor Melissa Freiberg and cohort partners Angela, Sommer, Stacey, Jessica, and Andrew have all provided valuable insight and feedback throughout our program. I would also like to thank my husband Christopher Bohlman, my parents Michael & Susan Meixner, and my son Joshua Bohlman for giving me the time to complete this Master's program and being patient with me.

Abstract

The purpose of this study was to determine if MAP (Measures of Academic Progress) testing from the NWEA (Northwest Evaluation Association) is a fair and appropriate test for English language learners. Ten elementary school students in a small suburban school were studied during the 2011-2012 school year. Their scores in reading and math in both fall and spring testing windows were compared with standardized growth targets. Their English language proficiency levels and classroom performance were also considered. The majority of students studied were able to reach recommended growth targets in reading and math. It is suggested that MAP is a fair test for these students, but is also recommended that teachers explicitly teach content and technical vocabulary, utilize appropriate accommodations, and become familiar with and use WIDA (World Class Instructional Design and Assessment) performance indicators for ELL students in order to enhance student performance on the test.

Table of Contents

Chapter	Page
Introduction	5
Literature Review	9
Overview of Required Testing	9
Testing Accommodations	13
Linguistic Complexity & Demands	15
Fairness & Validity	18
Measures of Academic Progress (MAP) Testing	19
Design & Methodology	22
Demographics	22
Testing	23
Results	25
MAP Testing Data	25
Language Proficiency & Classroom Performance	28
Teacher Qualifications	30
Conclusions	31
Analysis of Results	31
Implications for Classroom Instruction	32
Implications for School District	33
References	35
Appendices	37

Introduction

This project will examine the assessment scores of English Language Learners (ELLs), specifically on a district-mandated Measures of Academic Progress (MAP) test, to determine whether or not the test is appropriate for ELL students, and if not, to provide options to teachers to help improve scores. As a classroom teacher who is also certified to teach English as a Second Language (ESL), it is my goal to enhance my own teaching and practice to improve my students' performance on standardized assessments and to provide resources and assistance to other teachers who are also seeking to reach all students.

The issues of fairness and validity in assessment for ELL students have become an important part of my practice since I received my certification in ESL in 2008. I serve on the district ESL committee and work in a district that is still developing its ESL program. Although our ELL population is steadily growing, our staff is not. In the 2010 – 2011 school year, this district had a full-time ESL teacher as well as two bilingual educational assistants, one who spoke Spanish and the other Hmong. For the 2011 – 2012 school year, that full-time ESL teaching position was dropped to a 60% position. Our committee is looking for ways to incorporate ESL training and resources in schools in a way that does not add cost to programs. In order to do this and present information to the school board, it is important for us to have background information on second language acquisition and best practices for ELL students.

My district also utilizes the MAP (Measures of Academic Progress) test in the fall and spring of each year to chart student growth. This is a computer-based assessment that defines growth targets on an individualized scale for each student. The district uses this data to plan future instruction as well as for RtI (Response to Intervention) implementation. If the school

board and school district want to make Annual Yearly Progress (AYP), then it is vital that we are providing the optimum assessments that truly measure what our students know.

According to the National Education Association (NEA), “English Language Learners (ELLs) are the fastest growing segment of the public school population. Over the past 15 years, the number of ELL students has nearly doubled—to about 5 million. By 2015, ELL enrollment in U.S. schools will reach 10 million and, by 2025, nearly one out of every four public school students will be an English Language Learner” (National Education Association, 2008, online). In our current political climate of testing and accountability, and with the increasing numbers of ELL students in our Wisconsin public schools, it behooves educators to fight for basic fairness and validity in standardized tests. We frequently complain about the tests and the amount of time spent preparing students for them. However, I think we need to spend even more time, especially with our ELL students, preparing them for the academic language demands of the tests as long as they are required to take them. Otherwise, we do them a great disservice. It is a shame when a student’s intelligence and knowledge are unknown because of his/her inability to demonstrate them in a way that meets our standards. If ELL students are not performing well on the MAPS test, then it is up to educators in this district to either find a different test or develop ways to help the students be more successful.

President Barack Obama, in his plan for re-authorizing the Elementary & Secondary Education Act (ESEA), states, “every child in America deserves a world-class education” (Department of Education, 2010, online). Our country has long valued education and has worked hard to make schools fair and accessible to all students, from legislation on desegregation to Title IX. This must include children of illegal immigrants as well as all other

language minority students. The stand for a quality education for all is admirable; however, the plan itself is inadequate.

The current plan from the Department of Education will “provide significant formula grants to help states and school districts implement high-quality language instruction programs to improve the education of English Learners” (Dept. of Ed, 2010, online). It also requires each state to use a valid and reliable English language proficiency test to help determine placement for ELL students. Wisconsin is part of the WIDA (World-Class Instructional Design & Assessment) consortium, and so gives the ACCESS (Assessing Comprehension and Communication in English State-to-State) test to all ELL students. The ACCESS test is a test of language proficiency and tests ELL students’ abilities across four domains of language: reading, writing, speaking, and listening. According to WIDA, “The WIDA framework recognizes the continuum of language development within the four domains with six English language proficiency levels. These levels describe the spectrum of a learner’s progression from knowing little to no English to acquiring the English skills necessary to be successful in an English-only mainstream classroom without extra support. This final, exit stage for ELL status is designated Level 6 (formerly ELL)” (WIDA, 2007, online). “Formerly ELL” refers to a student who has enough English language proficiency to be comparable with his/her peers in a mainstream English classroom and does not need services anymore. Once English language proficiency (ELP) level is determined, WIDA also provides model academic standards for each grade level, subject area, and language proficiency level. These can be very helpful for teachers in determining what a particular student with a particular ELP is capable of in the classroom.

The Wisconsin Department of Public Instruction states that Wisconsin serves approximately 49,000 ELL students who speak more than 137 different native languages. The

two largest language groups are Spanish and Hmong. In the 2009-2010 school year, 74% of ELL students were required to achieve a proficient or advanced score on the state assessment in reading/language arts in order to meet AYP. 58% had to reach proficient or advanced standing in mathematics (Wisconsin Department of Public Instruction, 2010, online). The stated goal of the federal government for 100% of students to be proficient or advanced by 2014 applies to these students as well. With the move to tie teacher evaluations to test results, it becomes increasingly important to evaluate the validity and effectiveness of tests for various populations of students.

This integrative project will seek to describe the performance of second through fourth grade ELL students in a suburban Wisconsin school district on the MAPS test. It will detail their fall and spring scores, percentile rankings, and growth targets. My research will also disaggregate the scores according to ELP level as determined by ACCESS testing. ACCESS is given in December and January, meaning they are tested between the fall and spring MAP test.

Based on this information, I hope to make some conclusions about the appropriateness, equity, and validity of this test for our ELL students. If it is not considered a valid test for them, I will research available alternatives and make a recommendation to my district. If that is not possible, then I will determine, based on the information generated by the testing, effective ways for classroom teachers to help the students increase their performance and meet growth targets. If MAP is a valid test for ELL students, then I will offer ways for classroom teachers to increase the scores as well.

Literature Review

Overview of Required Testing

Several pieces of federal legislation affect standardized testing practices for ELL students. The first, the Elementary and Secondary Education Act (ESEA), was passed in 1965 and mandated funding of K-12 public education, including those students who were the neediest. Title VII of the ESEA specifically addresses bilingual education and requires schools to provide language support to ELL students. According to Kate Menken (2010), Title VII also “focused on creating opportunities for language learning that would result in equitable outcomes” (p. 122).

The No Child Left Behind Act (NCLB), passed in the United States Congress in 2001, was a re-authorization of the ESEA. NCLB requires that all students, including ELL students, make adequate yearly progress, as measured by standardized tests, and that all students are proficient in reading and mathematics by 2014 (Menken, 2010; DeVoe, 2007). Title VII of the ESEA was replaced with Title III, the English Language Acquisition, Language Enhancement, and Academic Achievement Act. As part of Title III, ELL students “must take tests of English language proficiency to measure their acquisition of English, and they must also take – and pass – the same tests of academic content as those taken by native English speakers” (Menken, 2010, p. 122). Since 2007, this has applied to all ELL students who have been in the United States for at least one year. Students who have been here for less than one year are exempt from the reading test but not the other content assessments (Wisconsin DPI, 2011).

Wisconsin has sought to conform to these mandates by having all ELL students take the ACCESS test, which tests English Language Proficiency (ELP). Unfortunately, ACCESS is not the only standardized test that ELLs are required to take. “LEP (Limited English Proficient)

students must be included in a State's assessment of academic achievement in reading/language arts and mathematics, and must receive appropriate accommodations and, to the extent practicable, native language assessments" (Office of Elementary and Secondary Education, 2006, online). All ELL students who have been in the United States for at least one year, also take the Wisconsin Knowledge and Concepts Exam (WKCE), given to all students, grades 3 – 12, at various times and testing various subjects. Again, students who have attended school for less than twelve months in the United States are exempt from this testing; however, it takes five to seven years for an ELL student to become academically competent in English (Wisconsin Center for Education and Research, 2007). According to Robert Linqanti, "because of the extended time required to learn sufficient English, ELLs are likely to be inaccurately assessed on their content knowledge as measured by standardized academic assessments" (cited in Policy Analysis, 2011, p. 16). Second language acquisition is also affected by aptitude, intelligence, personality, motivation, attitude, and the learner's preferences and beliefs, as well as the age at which the acquisition begins (Lightbown & Spada, 1999, p. 68). A twelve-month exemption from high-stakes testing is not enough to achieve the basic academic competence in a language needed to pass a content assessment.

According to John W. Young (2009), from Educational Testing Service, "it should be recognized that how well a student has performed on an ELP assessment will likely have an impact on his or her performance on content tests. In fact, standards-based content assessments generally assume something close to grade-level English language proficiency, so that the scores of students who are much below that level are likely to be invalid indicators of their content knowledge" (p. 123). A student who has been in the United States schools for two years is likely to have some basic interpersonal communication skill and is beginning to acquire some

content vocabulary. However, an ELL student in high school who must take the WKCE given to tenth graders each year, takes a content test written at a tenth grade level. Although he/she may be very intelligent in the content area if assessed in his/her native language, he/she may not be able to show his/her knowledge on a test given in English, a language he/she is only beginning to acquire (Policy Analysis, 2011). Unless the WKCE and other state mandated assessments can be given in the native language of the student, and therefore become a test of content rather than English proficiency, students should not be required to take these tests until they reach English language proficiency. From this information, we can judge that being in the country for twelve months is not long enough to attain grade level academic proficiency, even for a kindergartener.

In their book on ESL methodology, Suzanne Peregoy and Owen Boyle (2005) comment that, “because of the life long consequences of educational decisions based on high-stakes testing, it is essential that these tests be proven both fair and valid for all students...constant scrutiny is needed to monitor the effects of high-stakes testing to ensure that all students are provided meaningful and equitable access to a high-quality education” (p. 20). ELL students face tremendous difficulty when they encounter a content assessment. Willner, Rivera, and Acosta (2009) found that when ELL students take a content area test, they “may take longer to process the language of the test and frequently encounter difficulty accessing the content of test items due to unfamiliar language, cultural references, or format” (p. 697). Not all of these things can be mitigated through accommodation; in fact, some of them simply require a greater proficiency with English that can only be acquired through longer exposure. However, qualified teachers who are familiar with methods and best practices for teaching ELL students can use strategies and introduce accommodations to students that can make testing situations easier.

The Next Generation Assessment Task Force has recommended that the WKCE be phased out and replaced with a new statewide content assessment. The Task Force recognizes that, “Wisconsin educators are increasingly interested in receiving more frequent and more detailed data on the strengths and needs of their individual students” (Wisconsin DPI, August 2009, online). They advocate using assessments more in line with Measures of Academic Progress (MAP), which provide scores and information on progress almost immediately, and therefore can be used in the classroom.

Wisconsin is also part of a consortium of states that is said to be developing new assessments as part of the Common Core Standards Initiative. This initiative will include formative assessments in order to improve instruction and is hoped to increase scores on the summative assessments (Zehr, 2011; Wisconsin Center for Education Research, 2009). The FLARE (Formative Language Assessment Records for ELLs) project, commissioned by WIDA, is attempting to do just that – it is:

designed to assist teachers in measuring student progress in developing the essential language needed for success in academic classes. Quality, ongoing feedback from these assessments will help teachers and students target learning objectives and thus more effectively prepare students to accelerate learning toward the goal of language proficiency and post-secondary academic success (Wisconsin Center for Education Research, 2009, online).

Joining the WIDA consortium, as Wisconsin has, also provides continuity and consistency in educating our ELL students, as well as an accountability system accepted by the federal government’s Title III mandates.

Much of the research on the testing of ELL students looks at the variety and effectiveness of accommodations allowed by federal legislation. Researchers have also been interested in the linguistic complexity and demands of a content test written in grade-level English. Still others look at the fairness and validity issues surrounding these tests for students who may have been in the country for a very short time.

Testing Accommodations

Students who have been in the country for at least twelve months are required by Title III and other federal legislation to take the state tests in reading, language arts, science, and math (Abedi & Hejri, 2004; Willner et al., 2009; Menken, 2010). There has only been recent recognition that accommodations might be needed by ELL students to take these tests; so school and government officials are still working to define what is acceptable for these students. Too often, the list of accommodations for special education students is simply extended to ELL students without addressing their linguistic needs (Willner et al., 2009; Young et al., 2008).

As Willner et al. (2009) state, “effective ELL-responsive accommodations address the unique linguistic and socio-cultural needs of the student, but do not alter the construct being tested” (p. 697). Appropriate accommodations for ELL students include changes to the materials, procedure, or situation, such as presenting the test in the native language or via a qualified translator. Some students may access a bilingual dictionary or use extended time because of the longer time it takes to process information in a second language.

Abedi and Hejri (2004) studied the effectiveness of the various accommodations used by ELL students on the National Assessment of Educational Progress (NAEP). In their research, they first noted that accommodations were not assigned randomly to ELL students, making it

hard to differentiate between the initial differences the students had and the effectiveness of the accommodation itself. The NAEP data is problematic because of the small number of ELL students who received accommodations during the testing year (an average of about 5%). They had trouble in evaluating the validity of the NAEP accommodations, because in order to be valid, the accommodation should not affect the performance of non-ELL students. An experimental study was needed with random assignment to control and experimental groups using accommodations and not, to see whether the accommodation was valid (Abedi & Hejri, 2004, p. 374).

Despite these limitations, Abedi & Hejri decided to go ahead and study the performance of ELL students who received accommodations on the NAEP. They compared two separate assessments, from 1996 and 1998, and studied accommodated and non-accommodated students in grades 4 and 8 from each assessment. They matched students from one group with a student in the other group who met the same criteria: same test booklet, same school lunch program status, same Title I status, and same/similar parent education. They used a t-test to compare the performance of the two groups, expecting that the accommodated students would score higher than the non-accommodated students would, if the accommodation was effective. They also used a quasi-control group of non-accommodated non-ELL students to compare the accommodated students to as another measure of the effectiveness of the accommodations. The accommodations provided on the assessments included large print, extended time, bilingual dictionaries, test read aloud, small group, one on one, and use of a scribe or computer.

The researchers found that there was not a statistically significant difference between the mean scores in reading, writing, and civics among ELL students who received accommodations and those who did not. They concluded that providing these accommodations for ELL students

on a content-based standardized assessment did not reduce the performance gap between ELL students and native English speakers. They suggest that the accommodations available on the NAEP do not address the language needs of ELL students, and recommend that future NAEP assessments provide accommodation such as glossaries, dictionaries, and reducing the linguistic complexity of test items. Their study did focus on the collective group of accommodations, rather than each one, so it is possible that specific accommodations were more effective than others were. It would be vital for anyone conducting new research to differentiate between accommodations to see which ones made the most difference for ELL students. Indeed, as recommended by Willner et al. (2009), the accommodations must support the ELL student's linguistic needs. The student also needs to have been taught the content. Accommodations will not make a difference if the student has not been exposed to the material being tested. They also recommend that a team make the decisions on individual accommodations and to base them on the needs of the student, rather than assigning an accommodation to all ELL students. The final recommendation is to allow the students to practice using the accommodations before the actual test.

It is evident that accommodations may be helpful for ELL students when taking a content-based standardized assessment if the linguistic needs of the student are taken into account. The accommodation provided must be appropriate to the student and the assessment. Other accommodations, such as those provided to special education students, appear to make no difference in students' scores.

Linguistic Complexity and Demands

It takes an average of five to seven years for an ELL student to reach full English proficiency in academic content, assuming the best quality education and exposure to authentic

language use. For many students, especially those who come to the US as teenagers, this means that they are taking high-stakes content assessments when they are not yet fully proficient in English. (Young, 2009; Menken, 2010). The state tests are written in grade-level English and the complexity of the language, especially when the second language is dissimilar to English in structure and style, can be overwhelming for these students. Wolf and Leon (2009) note that “it raises a serious validity concern if test items contain unnecessary linguistic complexity that interferes with ELL students’ ability to show their content knowledge” (p. 140).

Wolf and Leon’s 2009 study investigated the language characteristics, differential-item functioning (DIF), and the relationship between them by studying multiple tests from several states in 2006. They specifically studied math and science assessments in grades 4 -8. The samples included all ELL students, high ELL students, and low ELL students (as rated by the states’ English language proficiency labels) and removed students with disabilities. They also included a reference group of non-ELL students. The population sizes of the samples were quite unbalanced, as the non-ELL students outnumbered the other groups in all three states. Their random sample of non-ELL students ended up being nine times larger than the other samples.

The following language characteristics were evaluated: academic vocabulary, grammatical features, sentence types, and cohesion. They also looked at more holistic features of the test such as form, visuals, and reliance on language. Differential Item Functioning (DIF) is “said to be present for an item when the probability of answering an item correctly is different between the two groups of interest who have the same ability level” (Wolf & Leon, 2009, p. 145). Through the process of rating the different test items, multiple raters coded each item using a specific protocol, and correlation analyses were conducted to determine the relationships

between the difficulty of the item, the language characteristics, and the DIF ratings. The number of occurrences of each characteristic per item was recorded.

The results showed that science items in general had higher means in linguistic complexity, and math items tended to have more visuals. The higher-grade levels showed a higher complexity of language. The category of reliance on language varied; on some tests, it was more crucial to understand the key vocabulary words in the items than to be able to process the whole sentence. On others, the reverse was true. They found more items exhibiting DIF for the low-ELL group than for the high-ELL group. They concluded that academic vocabulary was the most prominent linguistic characteristic, even on math tests, which are generally assumed to be less linguistically demanding. There was quite a bit of variance in linguistic complexity between the state tests, especially with regard to grammatical structure and discourse (p. 155).

Wolf and Leon (2009) suggest that the use of charts, graphs, and other visuals can help to lessen the impact of the language. They also suggest careful evaluation of the types of vocabulary used when writing test items. Some academic and technical vocabulary must be included as it is part of the content knowledge being tested. General-academic vocabulary tends to be the stumbling block for some students (i.e. Compare/contrast, sequence, paraphrase) – those words that are not specific to a discipline but are necessary for understanding what is being asked in the question. This leads to a conclusion about the importance of teaching students this general-academic vocabulary in the classroom before they encounter it on the test.

Careful consideration of the linguistic understandings required of students who are taking these content assessments, as well as purposeful instruction in general-academic vocabulary and test-taking strategies can make a content based assessment more accessible for an ELL student.

Combined with accommodations discussed previously, these are a vital part of assisting ELL students in achieving proficiency on grade-level content assessments.

Fairness and Validity

If we can make accommodations and reduce the language complexity, it seems that there should be no problem in having ELL students take the state assessments. John W. Young (2009) from Educational Testing Service states that, “to support claims that an assessment is valid and fair for ELLs, we must show that scores from the assessment have the same meanings and interpretations as for other examinee groups, such as native English speakers” (p. 125).

In their 2008 study, Young et al. investigated one state’s scores in math and science in grades 5 and 7 in the 2005-2006 school year. They viewed it as an exploratory study, without a specific hypothesis, to see if the same number of constructs were measured for all students and if the items were fair for ELL students, whether or not they used accommodations. They studied 30,000 students in each of four groups: non-ELL students, ELL students with no accommodations, ELL students with translated directions, and ELL students who could use glossaries and bilingual word lists and dictionaries.

For all the tests studied, the mean scores of ELL students, even those who used accommodations, were significantly lower than the mean scores of the non-ELL students. The accommodations seemed to make more of a difference as the students got older, perhaps due to their increased proficiency with using the provided materials. Young et al. concluded after their item and DIF analyses that there was only a small difference in the groups’ performance on individual items. They could not provide an explanation for why the mean scores were so different, and concluded that there could be real differences in content knowledge and mastery for the different groups.

If ELL students are expected to take content-based tests, then teachers must be sure that they have been taught the content. This requires well-trained ESL teachers who are familiar with methods for teaching content to ELL students in ways that are comprehensible and profitable for them. These teachers can do much to improve test performance for their English language learners by providing as much content exposure and knowledge as possible.

Measures of Academic Progress (MAP) Testing

According to the Northwest Evaluation Association (NWEA), the creators of the MAP test, “MAP assessments provide detailed, actionable data about where each child is on their unique learning path” (NWEA, 2011, online). The test was developed by educators, and is under continual evaluation by educators to ensure student interest and engagement. MAP is unique in that it adapts to the student taking the test – as the student records a correct response, the computer generates a more challenging item for the next question. If the response is incorrect, than an easier question is given. This design allows the test to target the student at his/her individual learning level, and give questions at which he/she can succeed.

Based on student scores in the fall, as compared with the norm group, a growth target is set. Teachers can access reports about students showing growth over time, as well as areas of strength and weakness within each subject area. Suggestions are offered to teachers for helping students attain the growth target.

Every test item on a MAP assessment corresponds to a value on the RIT Scale (for Rasch UnIT), so educators gain a deep understanding of what a student knows. RIT assigns a value of difficulty to each item, and with an equal interval measurement, so the difference between scores is the same regardless of whether a student is at the top, bottom, or middle of the scale. RIT measures understanding

regardless of grade level, so the information helps to track a student's progress from year to year (NWEA, 2011, online).

Each MAP test consists of approximately 40 – 50 questions in a subject area and is a computerized, multiple choice, un-timed test. Students use the mouse or keyboard to indicate their answer choice and then click a button to continue to the next item. They are unable to go back and change an answer once they have continued to the next question. (Cizek, 2010; Gierl & Alves, 2010). Once the test is completed, a student's RIT score is available, with RIT ranges for each subtest within that subject area.

Cizek's Buros Review (2010) indicates that, according to the manual available from NWEA, after test questions were developed, people from a variety of backgrounds tested them for bias. Cizek notes that it would be beneficial to have more information on the racial diversity of the field-testing samples, procedures for evaluating differential item functioning (DIF), and expanding the bias review to look more at native language, rather than simply cultural and ethnic bias. The norm group data was collected between 2001 – 2004 and included approximately 2.3 million students from 32 states. In his Buros Review, Cizek also states that, "proportions are indicated by state and student ethnicity, but basic information such as relative composition of males/females, public/private, urbanicity, special needs students, and so on is not provided" (2010, p. 4). Cizek's review indicates that MAP scores are reliable and valid, although he does specify that more descriptive information should be provided, including data on ELL students.

Gierl and Alves's Buros Review (2010) gives more information about the norm group. According to their research, the "ethnic characteristics of the samples were comparable to the American population for most ethnic groups" (p. 8), although they do not provide specific

information. They also conclude that the scores on MAP should be considered reliable, but the evidence for validity is limited.

The district in this study utilizes MAP testing as one measure of a student's performance. ELL students in grades 2 – 12 take the MAP test at least once each year, sometimes two or three times. The scores from the test are used primarily by teachers to help determine instruction and placement but the measured growth at the end of the year is also used in each school's Continuous Growth plan. Each grade level has a target percentage of students who should reach their growth target, and after spring MAP testing, the scores are analyzed to see whether the grade level or classroom met the percentage.

The MAP does not necessarily correlate with WKCE scores or other high stakes testing; rather, its purpose is to inform instruction, which should help to raise scores on state tests. However, if the MAP is not a valid or reliable test for ELL students, then it is not going to help teachers shape the instruction of these students. It is the goal of this research to determine whether or not the MAP is an effective method of assessing ELL students, and if not, to provide some alternatives.

Design and Methodology

Demographics

The district studied is a suburban/rural district in northern Dane County, Wisconsin. Approximately 15% of the students in the district are not Caucasian, with just under 4% of the total student population considered Limited English Proficient. The district employs a 60% ESL teacher, as well as 2 bilingual assistants, one who speaks Hmong and the other Spanish.

This study took place at a K-4 elementary school in the downtown area of this community. It served approximately 275 students in the 2011-2012 school year across 14 homerooms. Class sizes range from 15 – 23. Just over 5% of the students at the school are considered Limited English Proficient (LEP) students (approximately 14 students), and the school has the highest free and reduced lunch population in the district (33%). 60% of the students are in grades 2 – 4 and take the Measures of Academic (MAP) test each fall and spring. The school has made Adequate Yearly Progress (according to NCLB mandates) for each of the past 3 years in all areas, although there are too few ELL students to measure their progress as a sub-group.

Of the ELL students at the school, there are ten second, third, and fourth graders enrolled for the 2011-2012 school year. These students will take the MAP test in the fall and spring, the English language proficiency test (ACCESS) in the winter, and the third and fourth graders will take the WKCE in late fall. Only one teacher at the school has ESL certification. The ELL students are clustered in classrooms at each grade level and the bilingual assistants come regularly to work with students in those classrooms. The second through fourth grade ELL students at the school scored at levels 3.6 – 5.5 on the 2010 ACCESS test.

Testing

MAP and ACCESS test score data will be collected at three points during the school year, as well as anecdotal evidence regarding teacher preparation, perceptions, and services received by ELL students from teachers throughout the year. The scores of ELL students will be compared with the recommended growth targets provided by NWEA. One limitation of the study is the relatively small population at this elementary school. However, if significant issues are found with the testing, then the study could be expanded to the other elementary schools in the district and recommendations made to allow the teachers to assist their ELL students with the test.

All second through fourth grade students will take the MAP test in September 2011. Their scores in reading and mathematics will be collected. ELL students' scores in each grade level will be disaggregated according to language proficiency level (as determined by 2010 ACCESS scores) and native language. Along with the RIT (Rasch Unit) score in each subject area, students are also given a growth target relative to their initial score. Throughout the rest of the semester, the teachers who have these students in their classrooms will be surveyed to determine the types and amount of support given to their ELL students (see Appendix A).

In December 2011, ELL students will take the ACCESS test, which tests their language proficiency. The 2010 score and 2011 scores will be evaluated for any change. Teachers will again be surveyed regarding their familiarity with language proficiency levels and the WIDA standards (see Appendix B).

When students take the spring MAP test, the scores will again be compared to the fall score in both reading and math and disaggregated according to language proficiency (from

ACCESS 2011 scores) and native language. The scores will also be ordered based on reaching growth targets. The data that is of most interest to the district is the percentage of students who meet growth targets. This study will look at scores to see if there seems to be a relationship between meeting growth targets and proficiency levels, native language, classroom support, etc.

If this research shows that ELL students are meeting growth targets, then it may be determined that MAP can be an appropriate test for ELL students. If this research shows that success on the MAP test is inconsistent depending on language proficiency or native language, then suggestions will be offered for classroom teachers to help ELL students at varying proficiency levels be more successful at meeting growth targets. If the research shows that ELL students are consistently not meeting growth targets, then the MAP test itself must be further evaluated to determine its effectiveness for ELL students.

Results

MAP Testing Data

For the purposes of this study, ten ELL students were evaluated at an elementary school in this suburban district. There are four second graders, two third graders, and four fourth graders. Three of the second graders are in a multi-age first and second grade classroom. Six of the students are Spanish speakers, three speak Hmong, and one speaks Russian. The total number of ELL students at the school is 16, making them 5.7% (16 out of 280) of the total student population. For this study, students in kindergarten and first grade were not included since they do not take the MAP (Measures of Academic Progress) test. Of the total ELL population at the school, 56% speak Spanish (9 out of 16), 38% speak Hmong (6 out of 16), and 6% speak Russian (1 out of 16). It should also be noted that students B & G are siblings, as are students C & I.

According to the NWEA (Northwest Evaluation Association), specific RIT (Rasch UnIt) scores on the MAP test are expected in the fall and spring for each grade level in both mathematics and reading. Figure 1 gives the median expected fall scores in each grade level and subject, as well as a score requiring intervention (NWEA & Intervention Guidelines, 2009).

	Fall Median Math	Fall Intervention Math	Fall Median Reading	Fall Intervention Reading
2 nd grade	179	172	179	171
3 rd grade	192	185	192	183
4 th grade	203	196	201	193

Figure 1

Figure 2 provides fall MAP scores of each ELL student, in both subjects. (Scores requiring intervention are indicated with an asterisk.) Students took the tests during separate weeks in September 2011. It was an untimed test, and since none of these students has an Individualized Education Plan (IEP), no accommodations were made. All students, including native English speakers, are allowed to have a word pronounced for them on the mathematics test. The multiple-choice test was taken individually on the computer by each student. Score reports are provided to the district by NWEA upon completion of the school's testing window.

Student	Grade	Native Language	Fall Math	Math Percentile	Fall Reading	Reading Percentile
A	2**	Spanish	181	59%	184	70%
B	2**	Spanish	172*	32%	170*	35%
C	2**	Hmong	184	67%	175	48%
D	2	Spanish	171*	29%	165*	24%
E	3	Hmong	177*	11%	170*	9%
F	3	Russian	204	83%	197	68%
G	4	Spanish	205	54%	201	53%
H	4	Spanish	200	39%	194	34%
I	4	Hmong	198	33%	190*	25%
J	4	Spanish	216	82%	203	59%

Figure 2 ** denotes student is in a multi-age 1st/2nd grade class

NWEA also lists typical growth for each RIT score at each grade level. This is the amount of growth a student is expected to make over the course of the term. For the purposes of this project, the term was fall to spring, although typical growth can be generated for a fall-to-fall

term or a spring-to-spring term. Figure 3 lists each student's fall score, expected/typical growth, and spring score for each subject area. Students who did not meet growth targets are denoted with an asterisk.

Student	Fall Math	Typical Growth	Spring Math	Fall Reading	Typical Growth	Spring Reading
A	181	13 points	193*	184	13 points	198
B	172	14 points	193	170	14 points	188
C	184	12 points	191*	175	14 points	192
D	171	14 points	196	165	15 points	191
E	177	11 points	188	170	11 points	177 *
F	204	11 points	208*	197	9 points	203 *
G	205	9 points	220	201	7 points	213
H	200	9 points	214	194	7 points	203
I	198	8 points	206	190	8 points	206
J	216	9 points	228	203	7 points	212

Figure 3

The results of district MAP testing show that 70% of these ELL students met their math growth target and 80% met reading growth targets in the 2011 – 2012 school year. The results are inconclusive as to the influence of language ability on academic growth as measured by the MAP test, as some students with high ELP levels did not meet growth targets, and many with lower ELP levels did. Young (2009) reminds us “it should be recognized that how well a student has performed on an ELP assessment will likely have an impact on his or her performance on

content tests. (p. 123). These MAP tests are a measure of content knowledge, and with no accommodations given, the level of understanding of English must play a role in the success of the student on the test.

In the spring, the testing situation was unique in that the ELL students were tested in a small group to aid in data collection rather than being tested with their native English-speaking classmates. The time frame was within the four-week testing window, and no accommodations were made for spring testing. The third grade students both had a difficult time meeting growth targets. Student E was evaluated for special education over the course of the 2011-2012 school year but did not qualify. She received intensive one on one and small group instruction in both reading and math provided by the ELL support staff, reading specialists, and classroom teacher. She was able to meet the growth target in math but not reading. Student F is the other third grader, and she is the only Russian speaker in the school. Her ELP level is high, and she functions well in her classroom, but was unable to meet growth targets in either area. Her classroom teacher wishes to test her again with her classmates to see if she is able to meet growth targets. Student H was tested in reading and went home ill halfway through testing. She was then tested again with her classmates. This score is reflected in the above table.

Language Proficiency and Classroom Performance

All of these students also take the ACCESS test every December to measure their English language proficiency (ELP). The ELP levels range from 1 – 6, with a 6 being fully English proficient and means a student is exited from the formal program but will continue to be monitored. Figure 4 lists each student's ELP level from December 2010 and 2011 as well as their classroom performance in reading and math, according to their current teacher. Teachers were asked to rate their students in reading and math using the descriptors Poor, Fair, Average,

and Above Average (see Appendix A). Teachers' perceptions about student ability are mixed when compared with ELP level. In reading, within each grade level, the students with higher ELPs (2011) were perceived as performing better in the classroom. Math performance was rated similarly, though teachers rated three students differently in reading and math. Teacher perceptions are subjective and can be influenced by their own experiences, the performance of the class as a whole, and the variety of tasks performed in the classroom.

Student	ELP 2010	ELP 2011	Math Classroom Performance	Reading Classroom Performance
A	3.9	4.2	Average	Above Average
B	3.6	3.6	Average	Average
C	3.6	3.9	Average	Average
D	3.6	3.9	Fair	Average
E	3.7	3.8	Poor	Poor
F	4.7	5.2	Average	Average
G	4.3	4.7	Average	Average
H	5.4	4.4	Average	Fair
I	5.5	5.4	Average	Average
J	5.9	5.9	Average	Average

Figure 4

Most of the students receive some type of support from ESL staff in the district. Student D's parents refused any type of ESL support. Those who do receive support receive either in class or pull-out support from a Spanish speaking educational assistant, even if they do not speak

Spanish themselves. The range of support varies from simply sitting beside a student while they are working to working with students on class work in a separate room. Half of the students require translation on school materials that are sent home or interpreters for parent meetings.

Teacher Qualifications

Teachers report various experiences working with their ELL students. There were four teachers involved in this study, and all four have worked at this school for at least six years. One has ESL certification and the other three have all worked with ELL students in the past. They were given a survey (Appendix B) in January 2012 to ascertain their experience with and level of comfort working with ELLs, particularly in preparing them for standardized tests. All teachers reported that they do not feel equipped to help their ELLs do well on the MAP test. When asked what they would need, three out of four responded that they would like to know about possible modifications that could be made for ELL students on the test. One did not answer the question.

Conclusions

Analysis of Results

Since several students did not reach their expected growth targets despite the standard ESL programming support provided, the indication is that many ELL students may require more specific assistance with the testing process, particularly as the content on the test grows more complex. It does appear, however, that MAP is a fair and valid test for ELL students, as long as they have received instruction in content and technical vocabulary. They are able to meet stated growth targets in many cases, and their ability to meet the growth targets or not can be attributed to a variety of factors.

A factor that may also have influenced the results is that the researcher is one of the classroom teachers surveyed, and also has ESL certification. Four of the students are in the researcher's class and all four met their growth targets in both subject areas. It is possible that classroom instruction is better meeting the needs of ELL students in this classroom because of the teacher's training. It is also important to remember that several of the second grade students are in multi-age classrooms; a concept new to this district in this school year. It is unclear whether or not those classrooms performed lower in general than second graders did, however, the teachers are all new to teaching multi-age classrooms and report difficulties in managing both levels of students. In general, all second grade students are new to the concepts of MAP testing.

It is also noteworthy that all 1st and 2nd graders at this school had the opportunity to receive RtI (Response to Intervention) services in reading and literacy, while 3rd and 4th graders' RtI programming focused on mathematics. This could contribute to the second graders having more difficulty meeting growth targets in mathematics.

Implications for Classroom Instruction

As it becomes more and more common for classroom teachers to have student test results factor into their professional evaluations, teachers need to learn to be more effective in assisting all students with test preparation, as well as being aware of issues that may affect the students' performance on a test. If the MAP test is used as the main source of data, then students need to meet growth targets regardless of ELP level.

The following recommendations are given to teachers to help increase the number of students meeting MAP growth targets:

- 1) Teachers need to be informed about allowable accommodations for ELL students. The teachers surveyed were aware that they could provide accommodations for special education students, but not their ELL students. MAP tests are available in Spanish, if the student is able to read Spanish. Any modifications, such as bilingual dictionaries, should already be in use in the classroom, meaning teachers must be familiar with and comfortable using these tools with students before they take the test. When allowed, a bilingual assistant translating directions and test questions could be beneficial to students, particularly those with lower ELP levels.
- 2) Teachers need to use more appropriate and specific strategies for teaching ELL students in the classroom. Wolf and Leon (2009) also remind us of the importance of teaching academic vocabulary ahead of time in order to reduce the linguistic complexity for the student taking the test. In Young's 2008 study, we are also reminded that students need to have been exposed to the content in order to make gains. MAP provides detailed recommendations for each score range for how to work with students to get them to the next level. Teachers need to be sure that ELL students are exposed to the material, but

also that it is taught in a way that is understandable. Training in methods for teaching ELL students would be a valuable resource to teachers wishing to raise the test scores of their ELL students.

- 3) Teachers need to become familiar with the WIDA standards and performance indicators for each ELP level and language domain. The teachers surveyed indicated a knowledge that the standards existed, but no knowledge of how to use them to support their classroom instruction or to determine next steps for ELL students. Professional development for these teachers could be focused on the WIDA standards; unpacking them and designing lessons and/or classroom assessments more appropriate for ELL students.
- 4) Classroom teachers have noted that, in general, students who start out with higher RIT scores are faced with more difficult questions and technical vocabulary. Their expected growth, therefore, is generally lower than those with lower RIT scores in the fall. For ELL students who start with high RIT scores, a focus on technical vocabulary during classroom instruction is essential.

Implications for School District

It is clear that some ELL students are able to meet growth targets and make gains using the MAP test. Therefore, the test does seem to have validity for the ELL students in this study. As expected, students who are at higher language proficiency levels perform similarly to their native English-speaking peers, based on their ability to reach standard growth targets, if they have been exposed to the level of vocabulary expected by the test. MAP is a useful test for many students and teachers and provides specific areas of strength and weakness, individualized for each student. Its strength lies in its individual scoring and targets. Rather than assuming all

students will reach the same level, it allows for variance in student preparation and acknowledges progress that is made. Teachers use the data regularly in forming small intervention groups as well as identifying students who are excelling and providing enrichment opportunities for them. Teachers do not currently look at ELL students specifically as a group, but in the future, RtI groups could be targeted specifically around test vocabulary and preparation for ELL students. A focus on technical vocabulary and strategies such as compare/contrast, cause and effect, fact and opinion, etc. would benefit these students as much as exposure to specific content.

If ELL students continue taking the MAP test, the district must recognize that in order to assist teachers in helping ELL students, the district should make professional development a priority, and not an option. Much of the programming in the district focuses on students of lower proficiency levels, which is understandable considering their language needs. However, teachers of ELL students, particularly those students with ELP levels 4 and 5, must be given training in how to effectively prepare students in the academic vocabulary they will encounter on the test and be given ideas of accommodations and modifications they can make to both tests and classroom instruction. This can only be done with a commitment to a high level of quality ESL programming in the district, as well as adequate staff to assist in classroom instruction.

This school district, in its mission statement, seeks to give all students an excellent education. It is committed to engaging them, challenging them, and inspiring them to do their best. The teachers in the district are faced with a multitude of needs in their classrooms every day. With a commitment to excellence, and a mandate to help all students be successful, providing stronger instruction and support to ELL students should be a high priority.

Reference List

- Abedi, J. & Hejri, F. (2004). Accommodations for students with limited English proficiency in the national assessment of educational progress. *Applied Measurement in Education*, 17(4), 371-392.
- Cizek, G.J. (2010). Test review of Measures of Academic Progress. From R. A. Spies, J.F. Carlson, & K. F. Geisinger (Eds.), *The eighteenth mental measurements yearbook* [Electronic version]. Retrieved from the Buros Institute's Mental Measurement Yearbook Online database
- DeVoe, J.J. (2007). ELL testing: a state of flux. *District Administration*. Retrieved from <http://www.districtadministration.com/viewarticle.aspx?articleid=1293>
- Gierl, M.J. & Alves, C.B. (2010). Test review of Measures of Academic Progress. From R. A. Spies, J.F. Carlson, & K. F. Geisinger (Eds.), *The eighteenth mental measurements yearbook* [Electronic version]. Retrieved from the Buros Institute's Mental Measurement Yearbook Online database
- Lightbown, P.M. & Spada, Nina. (1999). *How Languages are Learned*. (2nd Ed.) Oxford: Oxford University Press.
- Menken, K. (2010). NCLB and English language learners: challenges and consequences. *Theory Into Practice*, 49:121-128.
- NEA Education Policy & Practice Department. (2008). *English language learners face unique challenges*. Retrieved from <http://www.weac.org/Libraries/PDF/ELL.sflb.ashx>
- Northwest Evaluation Association. (2011). *Measures of Academic Progress*. Retrieved from <http://www.nwea.org/products-services/computer-based-adaptive-assessments/map>
- Office of Elementary and Secondary Education. (September 2006). *Title I: improving the academic achievement of the disadvantaged*. Retrieved from <http://www2.ed.gov/legislation/FedRegister/finrule/2006-3/091306a.html>
- Peregoy, S. & Boyle, O. (2005). *Reading, Writing, and Learning in ESL: A Resource Book for K-12 Teachers*. (4th Ed). Boston: Pearson Education, Inc.
- Policy Analysis for California Education and Rennie Center for Education Research & Policy. (May 2011). *The Road Ahead for State Assessments*. MA: Rennie Center for Education Research & Policy.
- U.S. Department of Education. (May 2010). *Diverse learners*. Retrieved from <http://www2.ed.gov/policy/elsec/leg/blueprint/faq/diverse-learners.pdf>
- World Class Instructional Design & Assessment (WIDA) Consortium. (2007). *ACCESS for ELLs overview: essential background*. Retrieved from <http://www.wida.us/assessment/access/background.aspx>

- Willner, L.S., Rivera, C., & Acosta, B.D. (May 2009). Ensuring accommodations used in content assessments are responsive to English-language learners. *The Reading Teacher*. 62(8), 696-698.
- Wisconsin Center for Education Research. (2007). *WIDA: World-Class Instructional Design and Assessment*. Retrieved from <http://www.wida.us/>
- Wisconsin Center for Education Research. (2009). *Formative language assessment records for ELLs in secondary schools*. Retrieved from <http://flareassessment.org/>
- Wisconsin Department of Public Instruction. (February 2011). *Assessments for English language learners (ELLs)*. Retrieved from <http://dpi.state.wi.us/oea/ells.html>
- Wisconsin Department of Public Instruction. (September 2010). *Accountability for English Language Learners (ELLs)*. Retrieved from <http://www.dpi.state.wi.us/oea/ellamao.html>
- Wisconsin Department of Public Instruction. (January 2009). *Bilingual/ESL program*. Retrieved from <http://www.dpi.state.wi.us/ell/index.html>
- Wisconsin Department of Public Instruction. (August 2009). *Recommendations of the next generation assessment task force: crafting a balanced system of assessment in Wisconsin*. Retrieved from <http://www.dpi.state.wi.us/oea/pdf/NGTFbr.pdf>
- Wolf, M.K. & Leon, S. (2009). An investigation of the language demands in content assessments for English language learners. *Educational Assessment*. 14: 139-159.
- Young, J. (2009). A framework for test validity research on content assessments taken by English language learners. *Educational Assessment*. 14:122-138.
- Young, J., Cho, Y., Ling, G., Cline, F., Steinberg, J., & Stone, E. (2008). Validity and fairness of state standards-based assessments for English language learners. *Educational Assessment*. 13:170-192.
- Zehr, M. A. (2011). A vision for ELL inclusion in tests pegged to the 'common core'. *Education Week*. Retrieved from http://blogs.edweek.org/edweek/learning-the-language/2011/05/a_vision_for_ell_inclusion_in.html

Appendices

Appendix A

Staff Survey - ELL students and services.

Please circle your answer.

1. What grade do you teach? 2nd 3rd 4th
2. How many ELL students do you have in your classroom this year? 1 2 3 4 5

In thinking about your ELL students, please answer the following questions: (please use the same number consistently for each student)

What is the native language of this student?

Student 1:	Spanish	Hmong	Other: _____
Student 2:	Spanish	Hmong	Other: _____
Student 3:	Spanish	Hmong	Other: _____
Student 4:	Spanish	Hmong	Other: _____
Student 5:	Spanish	Hmong	Other: _____

Which students receive ESL services? (Please circle).

Student 1 Student 2 Student 3 Student 4 Student 5

Which students receive Tier 2 or 3 RtI services? (Please circle).

Student 1 Student 2 Student 3 Student 4 Student 5

Who provides ESL services to the students in your classroom? (Circle all that apply.)

ESL teacher (M. Leung) Spanish bilingual assistant Hmong bilingual assistant

Which students require an interpreter for conferences or communication with home?

Student 1 Student 2 Student 3 Student 4 Student 5

What is the average grade for each student for 1st quarter 2011-2012?

Student 1:	Reading _____	Math _____
Student 2:	Reading _____	Math _____
Student 3:	Reading _____	Math _____
Student 4:	Reading _____	Math _____

Student 5: Reading _____ Math _____

What was the student's RIT score for MAP in Fall 2011?

Student 1: Reading _____ Math _____

Student 2: Reading _____ Math _____

Student 3: Reading _____ Math _____

Student 4: Reading _____ Math _____

Student 5: Reading _____ Math _____

Please return completed survey to Rachel Bohlman. Your responses are anonymous and will be used in my research for my Capstone Project for my Master's Degree. If you have questions regarding this survey or my research, please contact me. Thank you for your time.

Appendix B

Staff Survey - Staff preparation

Please circle your answer.

1. What grade do you teach? 2nd 3rd 4th
2. How many ELL students do you have in your classroom this year? 1 2 3 4 5

How long have you been teaching? _____

Approximately how many of your years of teaching have you had ELL students in your class? ____

Which type of program have you most often experienced with your ELL students? (Circle one)

Pull-out In class support

Did you take any undergraduate coursework in ESL? Yes No

Have you taken any post-grade coursework in ESL? _____ If so, how many courses? _____

Have you had any professional development (not for credit) in ESL? Yes No

Are you familiar with the language proficiency levels provided by the ACCESS test and what they mean? Yes No

Are you familiar with the WIDA standards for ELL students? Yes No

How do you ensure that your ELL students understand the material you're teaching?

Do you feel like you have adequate training to help your ELLs do well on standardized tests?

Yes No

What information would help you better prepare your ELL students for standardized tests?

Please return completed survey to Rachel Bohlman. Your responses are anonymous and will be used in my research for my Capstone Project for my Master's Degree. If you have questions regarding this survey or my research, please contact me. Thank you for your time.