Chapter Two: Literature Review

Introduction

The inclusion of ELLs in high stakes standardized testing becomes problematic when one considers the fundamental function that language proficiency performs in the acquisition and assessment of content knowledge. Tests in content areas such as mathematics or science to some extent are also tests of language proficiency (Abedi, Courtney, Leon, Kao, & Azzam, 2006; Kieffer et al., 2009). This issue makes it difficult to make valid inferences about the achievement of ELLs. The fact that ELLs must take the state standardized mathematics test but may opt out of the English Language Arts test in the first year of attendance in U.S. schools is a reflection of the accepted notion that mathematics is less language dependent than other core areas. One tool for states to address this quandary is by providing accommodations to lessen the complexity of construct-irrelevant language, however, many of these accommodations tend to be ineffective or have questionable validity (Kieffer et al., 2009).

In order to better understand these issues, a review of the important literature relating to this topic will ensue following a discussion of the theoretical framework this research is based on. The literature falls into three general categories and will be examined in this order: 1) the language complexity of the math register, 2) the language impact on high stakes test performance of ELLs, and 3) accommodations used for ELLs to alleviate the language impact. Finally, a summary of the reviewed literature will be provided followed by a conclusion that identifies the gaps in the research that this study will address.

Theoretical Framework

The theoretical framework for this study is based on the socio-cultural theory of Vygotsky, more specifically his concept of the Zone of Proximal Development (ZPD). Vygotsky (1978), describes the ZPD as “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers" (pp. 86). In other words, learning occurs as a result of social interaction. The more ‘capable peer’ or adult serves as a bridge (or scaffold) between a learner’s current developmental level and their potential development. In this framework, language is not an isolated event that can be understood or acquired out of its social context.

Cummins (1981) distinction between basic interpersonal communicative skill (BICS) and cognitive academic language proficiency (CALP) relates to Vygotsky’s theory in that for the acquisition of CALP to occur it must be accessed within the learner’s zone of proximal development and scaffolded through social interaction. Cummins (2000) further elaborates on this concept:

> If students have not developed sufficient access to academic registers in either of their two languages, and if instruction does not provide the support that students need to develop this access, then their academic, linguistic, and cognitive development will not be stimulated through their classroom interactions (p. 106).

Through this quote Cummins clearly suggests that students require comprehensible input in order to further develop their learning. This concept of comprehensible input has given rise to a number of teaching and learning strategies for content area classrooms.

Despite the fact that this well founded theory is supported by empirical evidence, as shown below in the review of the literature, a number of researchers have found the theory to be controversial and have criticized Cummins' conversational and academic distinction. One of those critics, MacSwan (2000) argues that the notion of CALP represents a 'deficit theory.' This critique is founded in Chomsky's theoretical perspective that language development is complete before the age of five, thereby minimizing social and environmental influence and specifically the role that schools play in language development. MacSwan completely ignores the fact that our lexical knowledge continues to expand and that schools play a role in that expansion (Cummins, 2000).

Synthesis of Research Literature

The review of the important literature in this section will begin with a description of the language complexity of the mathematics register and then followed by an in depth look at studies examining the impact of language proficiency on high stakes test performance and the accommodations used to alleviate the language impact.

Language complexity of the mathematics register. Until recently, it was thought that mathematics learning is less language dependent than other core areas (Hansen-Thomas, 2009; Janzen, 2008; Lager, 2006) because it is more about numbers and symbols. Due to the state accountability mandates required for the sub-group of ELLs in the NCLB (2001) act, the attention of mathematics educators has started to shift toward the language needs of students in the math classroom (Lager, 2006). As a result, an increasing number of educators have focused on the content-specific vocabulary of mathematics; but since mathematics is a functional meaning making process (Lemke, 2003), more attention should be paid to engagement of students in the mathematics register through reading, writing and speaking rather than just rote memorization. This section will discuss the mathematics register, and some of the related language representations and features that have potential to confuse ELLs

The mathematics register. Cuevas (1984) defines the math register as "the meanings belonging to the natural language used in mathematics" (p.136). However, the language of mathematics contains more than just natural language (Martiniello, 2009), and this definition does not take into account that learning takes place socially and culturally (Moschkovich, 2007). O'Halloran (2000) expands this definition from the systemic linguistic functional perspective that "mathematical discourse is multisemiotic because it involves the use of the semiotic resources of mathematical symbolism, visual display and language" (p. 359). Lemke (2003) further described this perspective:

> Mathematics cannot be identified by the use of specialized mathematical symbolisms or any unique type of signs. Mathematics can be identified by the kinds of meanings it makes: meanings about addition, subtraction, multiplication, and division; about numerical difference and equality; about geometrical relationships of parallelism, orthogonality, similarity, congruence, tangency, etc., and many more in mathematical history. It is distinguished by these kinds of meanings, whether they are made by writing natural language, by drawing diagrams, or by formulating symbolic expressions (p. 1).

Learners actively negotiate these multiple meanings through a social and cultural lens as part of what this study considers as the mathematics register. Only a fraction of the mathematics register is included in the math items on standardized assessments, which currently do not take into account a socio-cultural perspective of learning (Solano-Flores & Trumbull, 2003). The learner must negotiate meaning among the different linguistic and non-linguistic representations in the context-reduced text without any form of feedback.

This process is further described by Martiniello (2009). First one must decode the problem's natural language which she refers to as "both the nonacademic or everyday language learned at home and other informal settings, and also the general cross-disciplinary academic language learned at school" (p. 162). Next, one must understand the content-specific terminology which includes vocabulary (i.e. *variable, sum, denominator)* and syntax (6 *divided by* 3, 6 *decreased by x is equal to 12*). By the same token, one must decode the nonlinguistic representations which include symbols (at times having their own syntax structures, i.e. $x > y - 8$), graphs, diagrams, tables and other visual representations.

Language features of mathematics that have the potential to confuse ELLs.

English language learners face the challenge of switching and translating between the mathematics register, the general English language register, and that of their native language to successfully solve mathematics problems (Lager, 2006). This challenge is further exacerbated by the fact that the algorithms and mathematical symbols used in their home country may differ from the ones used in U.S. schools (Wright & Li, 2008).

Brown, Cady, and Taylor (2009) cited many instances in mathematical language which may initially cause confusion for ELLs. A date in the U.S. is written as month/day/year (2/28/2011), but in Mexico it is written day/month/year

(28/2/2011). The majority of Europe and South America use a decimal comma to denote place value (13,8 would be read as thirteen and eight tenths) whereas North and Central America, Australia, and parts of Asia use a decimal to denote place value (13.8). Additionally, the U.S. is the only industrialized nation to use customary units of measurement (feet, yards, miles) as opposed to the metric system.

Researchers identified several difficulties ELLs may experience with mathematics textbooks and course materials. Mathematics textbooks commonly require both left-to-right and up-and-down eye movement, interpretation of non-linguistic representations, and must be read more slowly, perhaps multiple times, to develop comprehension (Wright & Li, 2008). Fillmore & Valdez (as cited in Lager, 2006) explain that ELLs may have trouble with written mathematics because meaning must be made from the language expressions, the order that they appear in and how they interconnect to be coherent.

Wright and Li (2008) conducted a linguistic analysis on released tests items from the 5^{th} grade mathematics Texas Assessment of Knowledge and Skills (TAKS) and student worksheets for two ELLs in a 5^{th} grade math class in Texas. The words from the assessment and the worksheets were compiled and compared using software and further analyses were conducted on both texts at the sentence level for syntactical complexity. As a result, Wright and Li (2008) found that the language complexity of the items on the 5^{th} grade Math Texas Assessment of Knowledge and Skills (TAKS) far exceeded that of the math worksheets provided to two Cambodian students.

Lager (2006), investigated the mathematics-language reading interactions that influence learning in algebra. The sample consisted of 221 sixth and eighth grade students, 133 of which were ELLs, who participated in the study by responding to nine items specifically regarding linear patterns. All of the items included text and visual patterns. Participants were asked to highlight any unknown or confusing words phrases. Cross tabulation of responses, content analyses of the work shown, and follow-up interviews with participants were also conducted to gain a better understanding of the issue. Results showed that there are language challenges in mathematics that can hinder both ELLs and non-ELLs.

Overall, non-ELLs performed better than ELLs with a mean difference of 1.3 or 0.6 standard deviations of difference in scores, yielding a medium effect size. Responses from over half of the participants that answered the first question incorrectly indicated that the most perplexing words were *previous, extension,* and *pattern.* All three words were considered troublesome for ELLs, while *extension* was the word of main concern for non-ELLs. Misunderstandings surrounding the polysemous phrase “Figure number (N)” caused incorrect responses from 25% of the participants regardless of language proficiency or grade level. The use of variables and parentheses were also cause for confusion. In addition to language challenges explicitly identified by students, it is equally important to examine the words that students chose not to highlight. After an analysis of their written responses and interview data there was evidence that students did not correctly comprehend some of the words and phrases that were not highlighted. Lager (2006) refers to this as “false knowing” and indicates concern that students are trying to develop a complex math register on a flawed foundation.

Research has also shown that a student's level of reading proficiency can be a strong indicator of mathematical success. Beal, Adams, and Cohen (2010) investigated the correlation between reading proficiency and mathematics achievement of 442 ninth grade Algebra 1 students in four Los Angeles, California high schools. Data sources included scores from state math tests and pre- and posttests developed by the researchers, progress reports from an online math tutorial program, self-report assessments on mathematics self-concept, and English conversational and reading proficiency levels for the 209 ELLs included in the study. The researchers found that math performance increased with English reading proficiency in a non-linear manner and that there may be a minimal reading level at which math performance will improve. Furthermore, they observed that reading proficiency was significantly related to math performance whereas speaking or listening proficiency was not.

Similarly, Lamb (2010) found that elementary and middle school students performed significantly worse on mathematics assessment items from the Texas Assessment of Knowledge and Skills (TAKS) that had a readability level that was above the student grade level.

As shown above, the complexity of the mathematics register can be problematic not only for ELLs but also non-ELLs. Yet, the situation is intensified for ELLs, as evidenced by the achievement gap in high stakes test scores for they must learn to translate between their native language, the general English register and the mathematics register. The research points to a need for more studies to focus on the facilitation of the mathematics register in the classroom, preparation for mathematics teachers regarding language, and the influence of previous schooling on the development of an ELL's mathematics register. The majority of these studies did not take into account the amount of schooling an ELL had in the U.S. The study by Wright and Li (2008) did address this issue but their research only included two ELLs from Cambodia which makes it difficult to generalize findings to a larger population. Furthermore, none of the studies take place in the Midwest which has seen recent surges in ELL populations. It is also imperative to look more specifically at how the linguistic complexity of items on high stakes mathematics tests relate to student performance.

Linguistic complexity and performance on math test items. In the past decade a number of researchers started to focus on the linguistic complexities of mathematics test items that could influence performance for ELLs (Abedi, Bailey, Butler, Castellon-Wellington, Leon, & Mirocha, 2005; Abedi & Lord, 2001; Lager, 2006; Martiniello, 2008, 2009; Ockey, 2007; Shaftel, Belton-Kocher, Glasnapp & Poggio, 2006; Wolf & Leon, 2009). They wanted to know if complexity or simplicity of language correlates with students' achievement on tests. Although there were differences among researchers, most of them found that language complexity strongly correlated with students' performance. However, the specific language features that influenced performance varied by test and grade level.

Ambiguous wording, item length, difficult vocabulary, syntactic complexity with longer sentences, and comparison problems are among some of the language features that affected performance (Shaftel et al., 2006). Most experts found that item length was the only language feature to have a consistent negative effect on item performance by ELLs (Martiniello, 2009). A handful of these studies and their results are described below.

Shaftel et al. (2006) analyzed 594 test items from the Kansas general mathematics assessment given at grades 4, 7, and 10 to examine the relationship between linguistic complexity and student performance. Through multiple regression analyses the researchers found that no pattern of item linguistic characteristics impacted item difficulties across groups. Difficult mathematics vocabulary had a consistent effect for all groups at every grade level; though, the researchers noted that this is a construct-relevant feature and should be expected to relate to item difficulty.

Likewise, two other studies conducted by Martiniello (2008, 2009) investigated the linguistic complexity of items that showed increased difficulty for ELLs as compared to Non-ELLs with equivalent math proficiency. Both studies utilized results from differential item functioning (DIF) procedures conducted on the 2003 Massachusetts Comprehensive Assessment System 4th grade mathematics test. Ten out of the 39 publicly released items from the MCAS showed DIF disfavoring ELLs (Martiniello, 2008, 2009). Think-aloud protocols administered to 24 ELL fourth graders in the 2008 study confirmed that the linguistic complexity of those 10 items was most likely the cause of DIF. Some of the shared linguistic features found in those 10 items include: multiple clauses, long noun phrases, unfamiliar vocabulary, polysemous words, and words or expressions referencing mainstream American culture (i.e. *coupon, spelling bee championship)* (pp. 358).

In addition to linguistic complexity, Martiniello (2008) found that the curriculum learning strand of Data Analysis, Statistics, and Probability was also a likely source of DIF disfavoring ELLs. Five out of the seven items assessing this strand were flagged as items that potentially disfavor ELLs. The researcher gives two possible explanations for this finding: either ELLs had less exposure to the curriculum content of this learning strand, or those items involved greater semantic and linguistic complexity. Nonetheless, she suggests that more research be conducted on the relationship of this strand and performance by ELLs.

Additionally, Martiniello (2009) explored the relationship between non-linguistic forms of representation (either schematic or pictorial) and DIF measures. Results showed that the impact of linguistic complexity was weakened when an item included a nonlinguistic schematic representation. Schematic representations were described as those representing spatial or mathematical relationships such as equations, diagrams, and tables. The researcher suggests that additional research should be conducted on the impact of non-linguistic representations on item performance as similar results may support the redesign of assessments to amend the linguistic complexity and reduce the need for accommodation.

Another study (Wolf & Leon, 2009) utilizing DIF procedures on math and science items from 11 assessments for Grades 4, 5, 7 and 8 taken from 3 different states found results consistent to the previous study. The relationship exhibited between the proportion of language in an item and performance suggested that the impact of language was “somewhat lessened” when charts, visuals, or graphs were included in the items (pp. 156).

Ockey (2007) also utilized a DIF procedure on 10 items from the NAEP math test (also used by Abedi & Lord, 2001), and found only 1 item exhibited DIF against

ELLs. Based on his results, Ockey concludes that ELLs can be tested validly on large scale standardized math assessments (pp. 161). However, this interpretation should be approached with caution because the item set only included 10 items out of the original corpus of 69 items. The ten items used for Ockey's study were subjectively chosen by researchers because they seemed to be the most likely to confuse a student due to language complexities, however, students may be confused more by items that were not included.

Although these studies focused on the correlation between language complexity and student achievement on math tests, none of them dealt with the issue of defining a "word problem," as it pertained to their study. Most use the term to loosely describe any mathematics item that includes words as opposed to an item composed strictly of numbers and symbols. This is evidenced by the following statement in Shaftel et al. (2006). "Furthermore, all items were presented as word problems, though the number of words per item ranged from 2 words (in six items at 4^{th} grade) to 177 words (in three items at 10^{th} grade), with a mean of 45 words (pp. 110)." Most studies do not even provide a definition or description of what they consider to be a word problem. This lack of definition complicates interpretations because the cognitive and language processes required for solving what is thought of as a traditional algebraic word problem are much more complex than that of a straightforward computational problem described in a sentence.

Moreover, studies should continue to focus on the effect that non-linguistic representations have on student performance as they may be a suitable construct relevant accommodation. Thus the specific types of accommodations that are provided for ELLs to alleviate language impact are posed.

Accommodations used for ELLs to alleviate the language impact. Appropriate test accommodations for ELLs are intended to reduce the impact of language proficiency on content assessments such as math without changing the target of assessment (Kieffer et al., 2009) and that do not give them advantage over students who do not receive the accommodation (Abedi et al., 2004). The use of such accommodations could allow those involved with high stakes testing to make more valid inferences about an ELL's knowledge, however, these accommodations must be effective and valid. In regard to accommodations, Kieffer et al. (2009), describe effectiveness as "the extent to which students receiving the accommodation demonstrate improved test scores," and validity as "the notion that the accommodation should improve the performance of students who require it but not affect the performance of students who do not" (pp. 1171).

Since inclusion of ELLs on standardized tests occurred only recently, the empirical research base on which accommodations are valid and effective is minimal. Two comprehensive overviews (Abedi et.al, 2006; Rivera & Collum, 2004) and one meta-analysis (Kieffer et al., 2009) of the research to date on accommodations were reviewed. The studies to be discussed were narrowed down to those that specifically dealt with ELLs and mathematics tests. Research on bilingual dictionaries and glossaries were excluded because they only studied the accommodation with science tests.

The most commonly used accommodations for ELLs can be divided into two categories: modifications to the test and to the test procedure. The empirical research in these two categories will be discussed followed by an account of the current state policies on accommodating ELLs.

Modifications to the test and to the test procedure. The majority of the research on accommodations include some type of modification to the language of the test (Abedi, Lord, & Plummer, 1997; Abedi et al., 2006; Abedi et al., 2004; Johnson & Monroe, 2004; Kieffer et al., 2009; Rivera & Collum, 2004; Robinson, 2010; Sato, Rabinowitz, Gallagher, & Huang, 2010). The most researched modifications include linguistic simplification, English dictionaries or glossaries, tests in the native language and extended time.

Linguistic modification. Research has shown that mathematics test items can be linguistically modified to simplify the English language load without changing the construct being measured (Abedi, et al., 2006; Abedi et. al, 2004; Abedi & Lord, 2001; Johnson & Monroe, 2004; Sato, et. al, 2010), although results of effectiveness vary. According to Abedi and Lord (2001), shortening sentences, removing low frequency construct irrelevant vocabulary, replacing conditional clauses with separate sentences, and changing complex question phrases to simple question words are among some of the ways an item could be modified or simplified.

Abedi & Lord (2001) gave a 25 question paper and pencil test containing 10 original National Assessment of Educational Progress (NAEP) test items, 10 linguistically modified questions and five noncomplex control items to 1, 174 eighth grade students. Thirty-one percent of the students were designated as ELLs. Results showed that ELLs scored lower on the math test than proficient English speakers but that linguistic modification of test items resulted in statistically significant differences, albeit a small fraction (.17) of an item difference.

Using a similar method, Sato and colleagues (2010) gave one of two 25 question math tests (one linguistically modified and one with original state and nationally released items) to 4,617 middle school students divided into three sub-groups of English proficiency. Differences in effect of linguistic modification for the three sub-groups differed by the scoring approach used, with only one out the four approaches showing any significant differences. In another study, Johnson and Monroe (2004) found that simplifying the language on test items for ELLs did not make a difference in their performance with findings indicated that ELLs performed slightly better on original test items.

English dictionaries and glossaries. In the meta-analysis conducted by Kieffer et al. (2009), the use of English language dictionaries or glossaries was the only accommodation found to have a statistically significant and positive average effect size. However, the types and descriptions of dictionaries and glossaries varied by study leaving no identifiable standard to determine the use of a dictionary or glossary as an appropriate form of accommodation (Rivera & Collum, 2004).

Abedi, Courtney & Leon (2003) compared four different accommodations for math assessments given to 4th and 8th graders. One of which was a customized English dictionary with a paper and pencil test, while another utilized pop-up glosses on a computer test. ELLs performance on the computerized test was significantly higher than other accommodations, conversely, interpretation of the results must take into account the factor of using a computer for testing. Another study by Abedi, Hofstetter, Baker & Lord (2001) resulted in ELLs benefiting most from the combination of an English glossary of non-technical terms and extra time, although non-ELLs also benefitted from this combination. Of further note, the use of glossary only was the least beneficial for ELLs.

A study conducted by Wolf, Kim, Kao, and Rivera (2009) after the meta-analysis took place found no significant difference in the performance of ELLs on a math test with a glossary provided. Further insight into the ineffectiveness of the glossary was garnered through student interviews, which indicated that the students did not use the glossary for various reasons including forgetfulness and inexperience using the tool.

Tests in the native language. Accommodations in the native language include written translation of test directions and/or items, oral repetition of the directions and/or items in the native language using audiotape or dvd (Rivera & Collum, 2004). The majority of the studies to date have focused only on Spanish translation of assessments (e.g. Hofstetter, 2003; Robinson, 2010; Solano-Flores & Li, 2009).

In Robinson's (2010) study, ELLs in kindergarten and 1st grade performed significantly better on mathematics assessments in Spanish rather than English. Hofstetter (2003) found that students who took a Spanish version of the NAEP mathematics assessment generally scored slightly lower than those taking the standard test. Yet he notes that students getting math instruction in Spanish performed better on the Spanish version than students who received math instruction in Spanish but took the standard test.

One unique study of note was by Solano-Flores and Li (2009). The researchers went a step further and included language variation in the first language (L1) to study the accommodation of test translation on NAEP mathematics items for different groups of students belonging to the broad linguistic group of Spanish speakers. Math items were given to students in bilingual education classes in a city on either the U.S East Coast or the U.S. West Coast in English, standard Spanish, and local Spanish (Mexican or Dominican). The results showed that the main source of score variation was the interaction of student, item and language (pp. 192).

Extra Time. A minimal amount of studies have exclusively focused on modifications to test procedures for ELLs, although these are among the most commonly permitted accommodations on national and state tests (Abedi et al., 2003; Abedi et al., 2004;). The accommodation of extra time is the easiest and cheapest to administer and it does not require the revision of any test items or directions (Abedi et al., 2004), however, the research is inconclusive on whether the strategy is valid (pp. 12). Both ELLs and English proficient students benefited from extra time on an assessment of NAEP math items in a study by Abedi, Lord, Hofstetter, and Baker (2000). In the 2003 study by Abedi et al., 4^{th} grade ELLs performed better on a mathematics assessment with extra time over standard conditions, use of customized dictionaries, or small group administration.

On the whole, the accommodations studied, although valid, have shown inconsistent and minimal if any effects on reducing the performance gap between ELLs and non-ELLs. Since linguistic modification of construct irrelevant language has not led to improved performance, more research should be conducted on alleviating the effects of the construct relevant language of the mathematics register without changing the validity and reliability, such as the use of non-linguistic representations mentioned earlier. More research in this area is especially important as interpretation of scores with and without accommodation is further complicated by the inconsistent policies and use of accommodations for ELLs by state.

Current state policies for accommodating ELLs. In a study of state policies for ELL accommodations Rivera and Collum (2004) found that most states had used a taxonomy developed for students with disabilities (SWD) to organize the 75 accommodations listed for ELLs. Forty-four of those accommodations were responsive to the needs of ELLs, while 31 were relevant only for SWDs (pp. 19). Including ELLs in the taxonomy for SWDs could obscure the differences in appropriateness of the accommodation for either sub-group of students. Furthermore, the policies vary in number, type, and use for certain content areas by state.

The State of Michigan currently offers a total of 63 accommodations organized in the taxonomy for SWDs (Office of Educational Assessment and Accountability [OEAA], 2009). Of the 63 accommodations, 23 are considered standard for ELLs and 9 of those are coded as *universal*, meaning any student is allowed to use those accommodations (i.e. use of a ruler supplied by the state, use of a highlighter).

Apart from the use of a word-for-word bilingual glossary, any accommodations providing direct linguistic support (i.e. directions or problems read aloud by test administrator) can only be used for ELLs that are: dominant in their native language and at a basic or lower intermediate level of English proficiency. In order to receive any direct linguistic support in the native language, the student must receive bilingual services at school to maintain their native language and be at a basic or lower intermediate level of English proficiency. There are no complete tests written in any languages other than English and there are no simplified English versions of the tests for ELLs.

Summary

Recognizing and understanding the relationship between language and mathematics learning for ELLs is crucial for educators and anyone associated with large scale assessment design and score interpretation. Incorporated with a socio-cultural perspective, the descriptions of the mathematics discourse by Lemke (2003) and O'Halloran (2000) constitute what is referred to in this study as a mathematics register. Learners must utilize the resources of mathematical symbolism, visual display, and natural language interdependently to negotiate mathematical meaning.

ELLs have the challenge of translating between three registers (native language, English, mathematics) to successfully solve math problems. This process includes decoding natural language, content-specific terminology (vocabulary and syntax), and non-linguistic representations. Furthermore, Brown, Cady, and Taylor (2009) identified several instances in the U.S. mathematics register that may cause confusion for ELLs including: the use of a decimal comma instead of a decimal point to indicate place value, the way that dates are written, and the use of customary units as opposed to metric units.

Another important observation concerning the challenges ELLs face with the mathematics register is that students in the study conducted by Lager (2006) had a sense of “false knowing” in regard to vocabulary in math problems. In other words, the students thought they understood more of the language than they actually did as indicated by their performance.

Research has also focused on the impact that construct-irrelevant language features of math items from high stakes tests has on ELL performance. The majority of these studies have found a relationship between language complexity and performance on math test items, but item length was the only language feature to consistently show a negative effect on performance (Martiniello, 2009).

Two studies conducted by Martiniello (2008, 2009) analyzed 10 items from the MCAS that showed DIF disfavoring ELLs. Further examination of these items confirmed that linguistic complexity (i.e. polysemous words, multiple clauses) could negatively influence the performance of an ELL. The researcher also found that half of the items showing DIF were from the learning strand Data Analysis, Statistics, and Probability. Results from the second study showed that the impact of linguistic complexity was weakened with the inclusion of a non-linguistic schematic representation. Martiniello suggested that further research be conducted on the relationship between item performance and both the Data Analysis, Statistics and Probability strand and non-linguistic representations.

Appropriate accommodations are intended to reduce the impact of construct irrelevant language features without disadvantaging students that do not receive those accommodations and without changing the target of assessment (Kieffer et al., 2009). One of the most studied accommodations is that of linguistic modification, or simplified English (Abedi, et al., 2006; Abedi et. al, 2004; Abedi & Lord, 2001;

Johnson & Monroe, 2004; Sato, et. al, 2010), but the overall outcome is inconclusive due to varying results of effectiveness. The meta-analysis conducted by Kieffer et al. (2009) determined that the use of English dictionaries or glossaries was the only accommodation studied to have a statistically significant and positive average affect size. Since the research regarding ELLs and high stakes testing is fairly recent, state policies on accommodation vary greatly in number, type and use for certain content areas (Rivera & Collum, 2004).

Conclusions

Valid inferences about ELLs performance on mathematics high stakes test are critical due to the consequences attached for ELLs and schools. The available research is minimal and for the most part inconclusive, perhaps because more of the issue is with the complexity of the construct-relevant language features. Further research into the effect of nonlinguistic representations on language impact is critical, as increased inclusion of such could act as a construct-relevant accommodation. Additionally, a comprehensive exploration of the relationship between item strand and performance is necessary. Any strand that consistently has a negative effect on performance should be examined for its accessibility to all students.

Since ELLs must acquire both the academic English register and the mathematics register simultaneously, teachers must be prepared to act as both a language and content area teacher to scaffold learning in the zone of proximal development. This has great implications for teacher preparation schools, professional development creators, and administrators. Finally, a consistent definition or description of what constitutes a ‘word problem’ must be provided and agreed upon by experts in the field in order for research results to be interpreted validly.