Validity of Accommodations for ELL students with Disabilities in Mathematics and English Proficiency Content Assessment

Jamal Abedi & Fereshteh Hejri
Advance Research & Data Analyses Center

Carol Lord
California State University, Long Beach

In collaboration with:

Kathryn Lee

August 2007

Submitted to:

Oklahoma State Department of Education

U.S. Department of Education

Council Chief State School Officers
The work reported herein was supported in part under the U.S. Department of Education Contract No. 84.368 as administered by the U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the position or policies of the National Center for Education Statistics or the U.S. Department of Education.
The generous efforts of several people informed the development of this study. We would like to thank all the individuals who participated in and informed this study. We are especially indebted to Van Anderson at the Oklahoma Department of Education for her support of this work. We are also grateful to Julia Lara, Don Long, Arthur Halbrook, Ellen Forte at the Council of Chief States School Officers and Guillermo Solano-Flores at the University of Colorado, Bolder, for their contribution to and support of this project. We are particularly thankful to Measurement Incorporated (MI, Michael Bunch and his staff) for providing free ELDA test booklets and for coding and scoring ELDA tests. They made a great contribution to the study. We also appreciate the assistance of:

Myra Chan  Seth Leon
Reshama Damle  Ebrahim Maddahian
Dan Farley  Shireen Pavri
Jose Garcia-Paine  Amy Poe
Allison Harris  Loretta Ransom
Cassandra Hawley  Karine Sarkissyan
Laura Kennedy  Hale Shelda
Melissa Kwon  Jisel Veg

We greatly appreciate the cooperation and assistance of the following states: California, Kentucky, New Mexico, Oklahoma, Oregon, and Virginia and all the test administrators, teachers, school staff, and students who helped us with this important work.
TABLE OF CONTENTS

Abstract 3
Introduction and Perspective 5
Review of related literature 10
  Assessment Accommodations 10
  ELL, SWD, and ELLWD Populations 11
  Types of Accommodations 13
  Effectiveness and Validity of Accommodations 16
  Accommodation by Translation 18
  Accommodation by Linguistic Modification 20
  Accommodation Providing Extra Time 21
Methodology 23
  Research Hypotheses/Questions 23
  Population and Sample 24
  Instruments 28
    Student, teacher, and school questionnaires 29
Research Design 30
Procedure 31
  Contacting the schools 33
  Preparing for testing 33
  Training test administrators 35
  Test administration 35
  Test scoring 36
Creation of the Linguistically Modified Mathematics Items 37
Item Ordering and Large Print Version 38
Bilingual Test Booklet 39
Accommodations: Effectiveness and Validity Issues 39
Examining the Validity of Accommodated Assessments 41
1. Performance Difference Approach 41
   2. Criterion-Related Validation Approach 42
   Examining the Effectiveness of Accommodations 43

Results 44
   Performance by Subgroups 48
   Effectiveness of Accommodations Used in This Study 54
   Validity of Assessment under the Accommodations Used in This Study 56

Discussion 66

Limitations 68

Appendices
   Appendix A: Student Background Questionnaire 88
   Appendix B: Teacher Questionnaire 91
   Appendix C: School Questionnaire 97
   Appendix D: Class Rosters 99
   Appendix E: State Approved Accommodations 104
   Appendix F: Parent Assent Letters 107
   Appendix G: Sample Test Administrator Training Agenda 112
   Appendix H: High School Participation Information Sheet 115
   Appendix I: Middle School Participation Information Sheet 118
   Appendix J: Elementary School Participation Information Sheet 121
   Appendix K: Math Test Administration Manuel 124
   Appendix L: All Math Tests Grade 5 135
   Appendix M: All Math Tests Grade 8 136
Validity of Accommodations for ELL students with Disabilities in Mathematics and English Proficiency Content Assessment

Abstract

This project addressed the needs of states to identify effective and valid accommodations for English language learners (ELLs) and students with disabilities (SWDs) in an effort to provide reliable and valid assessment English language proficiency as required by Title III and to fairly assess all students in mathematics as required by Title I of the No Child Left Behind (NCLB) Act. The project also provides information on the effectiveness and validity of accommodations for future national and state assessments for ELL students with disabilities (ELLWDs), a group of students that received limited attention yet faces the dual challenges of limited English proficiency and individual disabilities.

This study examined the validity of accommodations in two ways: 1) comparing the performance of accommodated non-ELL/non-SWD students with non-accommodated non-ELL/non-SWD students for whom accommodations are not intended, and 2) comparing the criterion-related validity of accommodated and non-accommodated assessment using structural equation modeling and multiple regression approaches.

A total number of 2,675 students were tested in the following six states: 1) California (Los Angeles Unified School District), 2) Kentucky, 3) New Mexico, 4) Oklahoma, 5) Oregon and 6) Virginia. Of this number, 555 (20.8%) students were in grade four, 521 (19.5%) in grade five, 458 (17.1%) in grade seven, 711 (26.6%) in grade eight, 232 (8.7%) in grade eleven and 196 (7.3%) in grade 12. Over forty seven percent of the students were female. Most of the students in the sample were recipients of free-reduced price lunch program (80.6%). Over 67% of the students were Latinos, 24.1% were Caucasian and 8.4% were African Americans. Of the total sample, 401 or 15% had some forms of disabilities. Due to our emphasis on the ELL population representation, over 47% of the sampled students were ELL students.

Students were assessed in mathematics and on the English Language Development Assessment (ELDA) using the reading and writing subscales. Listening and speaking subscales
of the ELDA were not included due to testing time limitations and logistical issues. In addition to mathematics and ELDA test data, background information from students, teachers and schools were collected. To examine the validity of accommodated assessment using external criteria, data were obtained on students’ test scores from state assessments through a Class Rosters prepared by teachers of selected classes. Background data from students, teachers and schools were used to control for sources of extraneous variables.

The results of the study shed light on the effectiveness of accommodations in reducing the performance gap between ELLs, SWDs, and ELLs with disabilities (ELLWD) on one hand and non-ELL/non-SWD students on the other hand. The data also helped establish the validity of the accommodated assessments as compared with the non-accommodated assessments. The results of this study suggested that most of the state-approved accommodations used in this study produced valid assessment outcomes. That is, these accommodations did not change the construct being measured. Therefore, findings of this study should be useful in state’s decisions concerning the continued use of accommodations from the state-approved list of accommodations for the three subgroups included in the study (ELLs, SWDs, and ELLWDs).

The results of effectiveness studies showed a trend of increased student performance under the accommodated assessments. However, due to the small sample size, the effectiveness of none of the accommodations used in this study reached a level of significance. Nonetheless, the results of our validity studies clearly suggested that the accommodations used did not compromise the validity of assessment. The results of multiple group confirmatory factor analyses as well as the findings of multiple regression using external criteria also confirmed that the provision of accommodation did not affect the validity of the assessments.

Our experience in conducting this study suggested that there are substantial challenges for comprehensive studies focusing on the assessment and accommodations for the three subgroups of special needs students (ELLs, SWDs, and ELLWDs) who lag far behind their peers. The goal of this study which included multiple sites from many locations nationwide was to have a nationally representative and large enough subjects to have power in testing effectiveness and validity of accommodations particularly for students in the low incident disability categories. However, in spite of the best efforts by the project team, it proved to be
extremely difficult to secure enough testing sites to provide sufficient data for a careful and comprehensive review of accommodation issues. Given the logistical difficulty, the outcome of this study provided evidence suggesting that many of the state-supported accommodations do not alter the construct being measured. Therefore, their use in the state assessment and accountability systems can be permitted. The findings of this study reveal the methodological issues in the assessment and accountability of special needs student population and point to the need for more attention to accommodation issues for these students.

The results of this study also reaffirmed the need of multiple indicators of student performance and the usefulness of obtaining relevant background variables. Such information helps to present a more comprehensive view of assessment and accountability requirements for these students.

**Introduction and Perspective**

To provide fair assessment for every child in the nation, both federal (e.g., NCLB Act of 2001, P.L. 107-110) and state legislation require the inclusion of all students, including English language learners (ELLs), students with disabilities (SWDs), and ELL students with disabilities (ELLWDs) in national and state assessments. Developing strategies for the inclusion of these subgroups is of utmost importance as ELL students continue to be a fast-growing population. According to a recent report by the U.S. Government Accountability Office, about 5 million ELL students were enrolled in schools, representing approximately 10 percent of all public school students (GAO, 2006). Between 1990 and 1997, the number of U.S. residents not born in the U.S. increased by 30%, from 19.8 million to 25.8 million (Hakuta & Beatty, 2000). Over 4.5 million ELL students were enrolled in public schools during the 2000-01 school year (Kindler, 2002). The rapid growth of this group demands that we consistently and accurately assess their content knowledge (in areas such as mathematics and science) as well as their level of English proficiency, and determine which students require English language services.

Schools are required to determine whether a child enrolling in school knows English well enough to succeed in an English-only classroom. Title VI of the Civil Rights Act, as interpreted by the U. S. Supreme Court in *Lau v. Nichols* (1974), prohibited discriminating against language-
minority children by ignoring their language related needs. For each child identified as ELL, the state must provide a comprehensible school curriculum that is comparable to programs for English-speaking students. The No Child Left Behind Act of 2001 (NCLB, Title III) requires ELL students to participate in a standardized measure of English language proficiency that documents development over time. According to NCLB (2002), assessments must provide for the inclusion of limited English proficient students, who shall be assessed in a valid and reliable manner and provided reasonable accommodations on assessments,… including, to the extent practicable, assessments in the language and form most likely to yield accurate data on what such students know and can do in academic content areas (NCLB, 2002).

The National Assessment of Educational Progress (NAEP) and assessments undertaken by individual states show ELL students with lower achievement levels than their peers in most states. The No Child Left Behind Act of 2001 requires states to ensure that all public school students meet standards of proficiency in mathematics and reading by 2014; NCLB designated ELLs as a distinct group for the reporting of state test results, and required that the ELL achievement gap be closed (NCLB, 2002).

Similarly, federal education policy requires that all students, including those with disabilities be included in national and state assessments. As mandated by NCLB and the Individuals with Disabilities Education Improvement Act (IDEIA, 2004), students with disabilities (including ELLs with disabilities) must take state regular assessments with or without accommodations or an alternate assessment. The alternate assessment can be based on grade-level achievement standards or on alternate standards for those students with the most significant cognitive disabilities, but no more than one percent of scores on alternate standards may be counted as proficient for NCLB accountability purposes. An additional two percent of scores may be counted for students with disabilities who can make significant progress toward grade-level content standards but may take longer to reach them than other students (Minnema, Thurlow, Anderson, & Stone, 2005).

Participation in state assessments may now be a legal requirement. However, in order to make valid inferences about student achievement and the effectiveness of schooling, more than mere participation is needed. Accommodations including modifications to the nature and
context of assessments have been devised to level the playing field.

Among the most important issues in the assessment of English language learners and students with disabilities is the concern over the validity of accommodations used for these students. If accommodations used for these students alter the construct being measured, then the outcomes from the accommodated and non-accommodated assessments may not be combined in analyses and reporting. Inaccuracies in the practices of accommodating ELLs, SWDs and ELLWDs may lead to the exclusion these students from the assessment and the accountability system, which may in turn have serious consequences for this group of students.

For ELLs, SWDs, and ELLWDs assessments using valid accommodations are likely to provide a more accurate picture of students’ content knowledge. The primary goal of this study is to provide fair opportunities for ELL students with disabilities to participate in the national and state assessments by identifying effective and valid accommodations for these students. Thus, the main focus of this study is students with the dual challenges of disability and limited English language proficiency. While there is some research on accommodations for SWDs, and on accommodations for ELLs, there is a striking lack of research on appropriate accommodations for ELLs with disabilities. Now that NCLB legislation requires the inclusion of all students in the national accountability picture, it is imperative to ask how all students can participate fairly in the assessment and accountability system. Two content areas for this study were selected, English proficiency and mathematics, both of which are of great importance in the national accountability requirements of the NCLB Act.

This study investigates the validity of accommodations for two purposes: 1) to provide data on the validity of accommodations for ELLs, SWDs, and ELLWDs that can inform future state and national assessments in mathematics and English proficiency as required by Title I and Title III of the NCLB Act, and 2) to assist with the implementation of English proficiency test for ELL students with disabilities.

Limited English proficiency and various forms of disabilities may prevent students from fully demonstrating their knowledge and capabilities. To level the playing field for these students, federal and state legislation requires testing accommodations—changes in the testing
process, in the test itself, or in the test response format—to accompany assessments. The goal of these accommodations is to provide a fair opportunity for these students to participate in the national and state assessments, without giving them an unintended advantage over non-ELL/non-SWD students who do not receive accommodations.

Accordingly, many states are using accommodations for these students (Abedi, Kim-Boscardin, & Larson, 2000; Rivera, Stansfield, Scialdone, & Sharkey, 2000; Thurlow & Bolt, 2001). The most commonly used accommodations for students with disabilities are Braille (allowed by 33 of the 48 states); computerized assessment (34 states); dictation of responses to a scribe (32 states); extended time (37 states); translation of instructions (40 states); allowance for marking answer in the test booklet (33 states); test items read aloud (34 states); read or simplified test directions (31 states); and test breaks (33 states) (Thurlow, et al., 2000). For ELL students, the common accommodations are: extended time (42 of the 48 states), use of glossary (26 states); use of an English dictionary (33 states); use of a bilingual dictionary (22 states); and linguistically-simplified test items (12 states). Rivera (2003) presents a list of commonly used accommodations for ELL students, a list that includes 73 accommodations.

However, the widespread use of these accommodations raises a number of issues and questions: Does using accommodations yield more valid inferences about student knowledge? Which students should be eligible for receiving accommodation(s), and what criteria should be used to decide their eligibility? What type of accommodation should be used? Are some accommodations more effective than others for all students or for particular groups of students? Do accommodations provide unfair advantage to the recipients? Is it meaningful to compare accommodated and non-accommodated scores? Without answers to these questions, the validity of assessment for students with limited English proficiency or physical or mental disabilities or both will be highly questionable.

The main thrust of this study was on accommodations that can be used in the assessment of ELL students’ level of English proficiency. There is an urgent need for reliable and valid English language proficiency measures to assess students’ level of English proficiency and to provide a valid classification of ELL students. Furthermore, Title III of the NCLB Act requires the assessment of ELL students’ English proficiency on an annual basis. Due to technical issues
with the existing English language proficiency tests, the U.S. Department of Education provided support through the Enhanced Assessment Instruments grants to states to develop reliable and valid measures of students’ proficiency in English. Several consortia of states were awarded grants to develop reliable and valid English proficiency tests. One group consists of eighteen states headed by the Nevada Department of Education and managed through the LEP State Collaborative on Assessment and Student Standards (LEP-SCASS) at the Council of Chief State School Officers (CCSSO). This consortium has developed the English Language Development Assessment (ELDA) for grades K-12 in four grade clusters (K-2, 3-5, 6-8, and 9-12) that measures English proficiency in four domains: 1) reading, 2) writing, 3) listening, and 4) speaking.

ELDA is a newly-developed assessment; one of the first assessments aligned to state English Language Proficiency (ELP) standards intended to measure the annual growth of a student’s English proficiency. Due to its short history, intensive validity and reliability analyses needed to be performed in order for it to properly serve states, districts, and students. Among the major issues that participating states face were those concerning the use of accommodations—whether or not to use accommodations in this assessment for some students and whether the accommodated and non-accommodated results can be combined for analyses and reporting purposes.

Included in the assessment are a group of ELL students with various forms of disabilities. For a valid assessment of English proficiency of this group of students, some forms of accommodations may be needed. For example, providing large print for students with visual impairment or hearing devices for students with hearing impairments may be necessary to level the playing field for these students. However, the main concern in providing such accommodations for ELL students with disabilities in the ELDA assessment is the validity of the accommodations used for these students. Do the accommodations alter the construct being measured (English language proficiency in this case)? Can the accommodated ELDA assessment results be combined with the non-accommodated assessments? This study intends to shed light on the validity of accommodations for ELL students with disabilities to inform the ELDA test development process in creating a more reliable and valid measure of English proficiency.
Review of related literature

Various types of accommodations have been proposed and used with ELL and SWD populations. A limited number of empirical studies have addressed the issues of effectiveness and validity of accommodations for these students. This chapter addresses the rationale for accommodated testing; the nature of ELL, SWD and ELLWD populations; legal mandates for assessment; types of accommodation and their effectiveness and validity; and implications for research.

Assessment Accommodations

According to the Individuals with Disabilities Education Act (IDEA) of 1997, students with disabilities are expected to participate in state and district assessments. Large-scale assessments are based on the assumption that they provide comparable test scores with presumably standardized testing conditions. However, some students with disabilities may be better able to demonstrate their knowledge if they are allowed to participate under accommodated testing conditions. The accommodations may produce a more accurate measure of the student’s knowledge. An effective accommodation will address the nature of the student’s disability, resulting in a more accurate score without affecting the test performance of non-disabled students; that is, the accommodation is not related to the construct that the test is designed to assess.

Butler and Stevens (1997, p. 5) view accommodations for ELL students as “support provided to students for a given testing event either through modification of the test itself or through modification of the testing procedure to help students access the content in English and better demonstrate what they know.” Or, as defined by the Center for Equity and Excellence in Education, George Washington University (2005), accommodations are: “any change to a test or testing situation that addresses a unique need of the student but does not alter the construct being measured.”

In some state programs a distinction is made between accommodation and modification: an accommodation does not alter the construct being measured, but a modification is a change that alters the construct (Sireci, Scarpati, & Li, 2005). Alternate assessments are tools used to
evaluate the performance of students who are unable to participate in state assessments even with accommodations (Thurlow, Elliott, & Ysseldyke, 2003). Most states have a policy or guideline stating that accommodations used in state testing must be same type used for instruction (Lazarus, Thurlow, Lail, Eisenbraun, & Kato, 2006). However, there may be a lack of alignment between specifications in IEP/504 accommodation plans and implementation during state assessments, classroom assessments and instruction (Shriner and DeStefano, 2003).

**ELL, SWD, and ELLWD Populations**

The term “students with disabilities” (SWD) is typically used to describe those students who have an Individualized Education Plan (IEP), as required by the 1975 Individuals with Disabilities Education Act (IDEA), or a 504 plan. “English language learners” (ELL) may be defined as students with non-English language backgrounds, with varying degrees of English proficiency, who have been identified as ELL or LEP (Limited English Proficient) according to state definitions (Minnema, Thurlow, Anderson, & Stone, 2005). As defined by the U.S. Department of Education, they are national-origin-minority students who are limited-English-proficient. However, the use of the term ELL rather than LEP has come to be used more often because of its focus on accomplishment rather than deficiency (August & Hakuta, 1997; Butler & Stevens, 1997; LaCelle-Peterson & Rivera, 1994).

Eighty percent of ELL students speak Spanish as their first language (U.S. Department of Education, 2007). Students classified as ELL or LEP may possess a range of individual characteristics in terms of language and cultural background, prior schooling outside the U.S., and length of time in the U.S. (Butler & Stevens, 1997; Rivera, Collum, Shafer Willner, & Sia Jr., 2006). Since 1979 the percentage of children speaking English with difficulty has nearly doubled (Federal Interagency Forum on Child and Family Statistics, 2005). According to the 2005 American Community Survey, 2.7 million students spoke a language other than English at home and reported speaking English less than “very well,” an increase of 1 million students over the 1980 census total (Fry, 2007). It is projected that 30 percent of the school-aged population in the U.S. will be ELLs by 2015 (Francis, Lesaux, Rivera, Kieffer, & Rivera, 2006). The number of ELLs in our schools will likely continue to grow if the population of school-age children who have immigrant parents increases as projected from 12 million in 2005 to 18 million in 2025.
As the total number of ELLs in our schools increases, the number of ELLs with disabilities can likewise be expected to increase, along with the need to understand which accommodations are effective and valid for assessments with these students (McCardle, Mele-McCarthy, & Leos, 2005; McCardle, Mele-McCarthy, Cutting, et al., 2005).

In the United States, more than 4 million students between the ages of 6 and 21 have at least one disability, as reported by Olson and Goldstein (1997). Students with learning disabilities (LD) account for over half of the population with disabilities (Artiles & Klingner, 2006). Based on responses to a written survey to local education agencies, the number of ELLs with disabilities (ELLWD) nationwide was reported as 357,325 (Zehler, Fleischman, Hopstock, Pendzick, & Stephenson, 2003). Minnema, Thurlow, VanGetson, and Jimenez (2005) note that approximately 11% of native English-speaking students have disabilities, but that the proportion of ELLs receiving special education services has been estimated as between 7.87% and 9% (D’Emilio, 2003; Zehler et al., 2003). This would indicate that ELLs are underrepresented as SWD.

Minnema, Thurlow, VanGetson, and Jimenez (2005) cite additional data suggesting under representation of ELLs in the SWD classification. In the 1996 NAEP science assessment, about 1% of the students were identified as ELLs with disabilities. But this number would be expected to be higher, in light of overall estimates of 12.7% ELLs and 10.7% SWD nationwide according to Minnema, Thurlow, Anderson, and Stone (2005). If the percent of ELLs with disabilities were the same as the percent of non-ELLs, the number of ELLs with disabilities would be approximately 1.36%. A study by Albus and Thurlow (2005) observes that ELLs with disabilities are not defined the same way across states and may be either over or under identified from state to state.

Other studies indicate that English language learners are overrepresented in special education programs (Artiles & Trent, 1994; Artiles, Trent & Palmer, 2004). One study found that students identified as having limited proficiency in both their first language (L1) and in English (L2) were consistently overrepresented in learning disabilities and language and speech disabilities classes and were more likely to be placed in special education programs than other ELL students (Artiles, Rueda, Salazar, & Higareda, 2005).
The research on ELL students in special education programs is limited (Donovan & Cross, 2002; Heller, Holtzman, & Messick, 1982). Schools face the difficult question of whether a student’s low performance is due to a disability or to language acquisition issues or a combination of both (Abedi, 2006; Klingner, Artiles, & Barleta, 2006). Language acquisition processes may be misunderstood as learning disabilities (Artiles & Ortiz, 2002; Collier & Hoover, 1987; Langdon, 1989; Ortiz, 1997; Rueda & Windmueller, 2006; Zehler & Fleischman, 2003) resulting in misclassification of ELL students as learning-disabled. Teachers may be unprepared to work effectively with ELL students and may be uncertain on how to adapt curriculum, instruction, and assessments to meet these students’ learning needs; this may result in a referral for special education services. A high percentage of referred students are found to qualify for special education services; possible factors include assessor bias and lack of adequate bilingual assessments and trained personnel.

States face complex challenges in striving to include all students in statewide testing programs. The inclusion of all students has generally not been considered in the construction of large-scale assessment instruments (Thompson, Johnstone, & Thurlow, 2002). States have demonstrated progress in participation and performance for students with disabilities and English learners (Thompson & Thurlow, 2003). The inclusion of ELLs with disabilities is doubly complex, since issues of both English language proficiency and disability must be considered (Minnema, Thurlow, Anderson, & Stone, 2005).

**Types of Accommodations**

Considerable variability has been found among states’ guidelines on accommodations in assessments such as differences in the terminology used; references to the concept of accommodation may include terms such as modification, mediation, alteration and adaptation. Studies of accommodations for students with disabilities include: Koenig and Bachman (2004); Mazzeo, Carlson, Voelkl, & Lutkus (2000); Thurlow, McGrew, Tindal, Thompson, Ysseldyke, and Elliott (2000), Thompson, Blount, and Thurlow (2002); Tindal, Anderson, Helwig, Miller, and Glasgow (2000); and Tindal and Fuchs (2000).

Types of accommodations have been classified in various ways. In 1996, the National
Center on Educational Outcomes (Elliott, Thurlow, & Ysseldyke, 1996) provided a four-way categorization of accommodations for students with disabilities. The categories were 1) Presentation, 2) Student Response, 3) Test Setting, and 4) Timing or Scheduling. Specific examples of accommodations in each category included the following:

1) Presentation:

✓ Increase size of answer bubbles
✓ Highlight key words or phrases in directions
✓ Read directions to student
✓ Provide audiotaped administration of sections
✓ Read questions aloud to a student

2) Student Response:

✓ Provide wider lines and/or wider margins
✓ Allow student to mark responses in booklet rather than on answer sheet
✓ Provide word processor

3) Test Setting:

✓ Provide special lighting
✓ Provide adaptive or special furniture
✓ Allow testing in a small group, study carrel, or individually

4) Timing or Scheduling:

✓ Extend the time allotted to complete the test
✓ Administer the test in several sessions, specify duration
✓ Allow frequent breaks during testing

Nearly a decade later, the findings of another study from the National Center on Educational Outcomes showed the increased attention to the inclusion in assessments of students with disabilities and the increase in provision of accommodations for them. An additional
category-equipment and materials accommodations-was included as well as new specific forms of accommodation (Lazarus, Thurlow, Lail, Eisenbraun, & Kato, 2006).

In the category of presentation accommodations, Lazarus, Thurlow, Lail, Eisenbraun, and Kato (2006) found that 48 states allowed large print, 44 allowed Braille, and 41 allowed directions to be read aloud to the student. Forty-three states allowed the directions portion of the assessment to be presented to the student via sign language; 38 allowed questions to be read by the examiner and clarified through restatement; 25 allowed visual cues such as arrows or stickers; 38 allowed someone other than the regular test examiner to administer the test. Other accommodations were reported with lesser frequency; for example 18 states allowed for audiotape presentation of test directions.

Among equipment and material accommodations, 42 states allowed magnification equipment, 39 allowed amplification equipment (e.g., hearing aids), 33 states allowed changes in the amount or placement of lighting or acoustics, 38 allowed place markers, 31 allowed noise buffers such as ear mufflers, and 32 allowed special furniture. The use of a calculator was allowed in some states, but allowed with restrictions in others. A number of response accommodations were allowed, but typically with some restrictions. These included use of a scribe, a computer, tape recorder, communication device, Brailler, or sign language response. Six states allowed a thesaurus. Changes in scheduling or timing of the assessment included extended time, allowed by 39 states and with restrictions in 6 other states. Many states allowed changes in the setting or location; 45 states allowed individual assessment and small group assessment.

Assessment accommodation policies for ELLs have varied among the states. The most common accommodations have been extra testing time and small group administration, individual administration, testing in a separate location or carrel or with more breaks, and the use of a bilingual dictionary or word list (Abedi, Hofstetter, & Lord, 2004). Less frequent accommodations have been translation of the test into the student’s native language, a bilingual version of the test, and a modified or simplified English version.

A study of state accommodation policies for ELLs identified 75 accommodation types,
with 44 of those addressing the linguistic needs of ELLs and no clear criteria for choosing accommodations (Rivera, 2006; Rivera, Collum, Shafer Willner, & Sia, 2006). Policies often provide a list of accommodations intended to address both ELLs and students with disabilities. An accommodation taxonomy more responsive to the needs of ELLs would include categories for 1) direct linguistic support and 2) indirect linguistic support. Accommodations providing direct linguistic support would involve adjustments to the text of the assessment with the intent of reducing the linguistic load; indirect linguistic support would involve adjustments to the conditions under which a test is taken to allow ELLs to more efficiently use their linguistic resources. Rivera (2006) recommends that states clearly distinguish the needs of ELLs from those of students with disabilities; for ELLs, the accommodation should be effective in supporting access to the content of the test.

**Effectiveness and Validity of Accommodations**

An accommodation is effective if it leads to improved test scores for the students who receive the accommodation. However, to be valid, an accommodation should be differentially effective. An accommodation may result in higher scores for all students. Although that may be considered a desirable outcome, it indicates that the accommodation may alter the construct being measured; then the accommodation would not be considered to be valid. Using accommodations without evidence of their effectiveness or validity is unproductive and unfair to students. Since empirical evidence on effectiveness and validity of various accommodations is limited, the current study seeks to address the need for better information so that more informed choices among accommodation options can be made.

Researchers have recently begun to study inclusion of ELLs with disabilities in large-scale assessments. In a study of the use of accommodations, the majority of respondents to a survey reported that ELLs with disabilities “always” (42%) or “usually” (52%) take the state tests, and “always” (30%) or “usually” (55%) use some form of accommodation (Minnema, Thurlow, VanGetson, & Jimenez, 2006b). Accommodations are more likely to be designed for special education than for ELLs. Of the respondents, 69% said that only special education accommodations were “usually” or “always” used. In another survey 62% of respondents said ELLs with disabilities use only special education accommodations “usually” or “always”
A recent review of the literature found only a few studies that specifically addressed accommodations used during testing for ELLs with disabilities (Abedi, 2007; Minnema, Thurlow, Anderson, & Stone, 2005). However, there has been much discussion of accommodations for ELLs, as well as accommodations for students with disabilities. In a review of research on accommodation use for students with disabilities, Tindal and Fuchs (2000) note the lack of consensus among states regarding appropriate accommodations. Comparisons of effectiveness and validity are difficult because of the lack of comparability between states, and without clear information about validity, decisions for individual students may be formulated idiosyncratically by IEP teams. Tindal, Heath, Hollenbeck, Almond, & Harniss (1998) review a number of studies in which various forms of accommodation were utilized. Accommodations used on large-scale tests are discussed.

Numerous studies have addressed the use of various forms of assessment accommodations for ELL students (Abedi, 1999; Abedi, 2004; Abedi, 2005; Abedi, 2006; Abedi, Courtney, & Leon, 2003; Abedi, Hofstetter, & Lord, 2004; Abedi & Lord, 2001; Abedi, Lord, & Hofstetter, 1998; Abedi, Lord, Hofstetter, & Baker, 2000; Abedi, Lord, & Plummer, 1997; Butler & Stevens, 1997; Francis, Lesaux, Rivrea, Kieffer, & Rivera, 2006; Hakuta & Beatty, 2000; Hambleton & Patsula, 1998; Kopriva, 2000; Olson & Goldstein, 1997; Rivera, Stansfield, Scialdone, & Sharkey, 2000; Shepard, Taylor, & Betebenner, 1998; Thurlow et al., 2000; Sireci, Li, & Scarpati, 2003; Solano-Flores & Li, 2006; Stansfield & Rivera, 2001; Zehler, Hopstock, Fleischman, & Greniuk, 1994). For ELLs the nature of the difficulty, the barrier to success on assessments, is linguistic. Of the many forms of accommodation proposed and in use, some address the language barrier directly (e.g. providing English or bilingual glossaries or by modifying the English to make the vocabulary and sentence structures easier to understand). The language barrier is addressed indirectly by other forms of accommodation, such as providing extra time for the student to comprehend and respond. These options are discussed in the following sections.
Accommodation by Translation

For ELLs, a plausible way to overcome the language barrier is to assess students in their native language. Research indicates that this approach to accommodation is useful only when students can demonstrate their content knowledge more effectively in their native language. This may be the case if the students have studied content area subject matter in their native language. However, if a student’s exposure to content and vocabulary (e.g., in mathematics or science) has been in English, the student may not be familiar with academic content-area vocabulary in the native language, and a native-language test would not be an effective accommodation (Abedi, Hofstetter & Lord, 2004; Butler & Stevens, 1997). An experimental study by Abedi, Lord, and Hofstetter (1998) found that eighth grade students who had received mathematics instruction in Spanish scored higher on NAEP mathematics items in Spanish than their peers did with similar items in English. Additionally, Hispanic ELLs who received classroom instruction in English scored higher on the mathematics test in English than their peers did on a test with corresponding items in Spanish.

When a test is translated from one language to another, or when test items are developed in parallel for more than one language, test validity may be compromised because of the difficulty of maintaining construct equivalence (American Institutes for Research, 1999; Kopriva, 2000; Sireci, 1997). Even with international guidelines for test translation and adaptation, translated assessments are technically difficult and time consuming and expensive to develop (Hambleton, 1994). A word or phrase in one language may not have an exact equivalent in another language; the word may occur frequently in one language, but its equivalent may be rare in the other language, adding to the difficulty of the item (Hambleton & Patsula, 1998). If a language has several dialects, a translation into one dialect may not be suitable for speakers of a different dialect (Olson & Goldstein, 1997).

One method to ensure accuracy of translations is called “back translation” (Brislin, 1970, 1986). This method is used to increase the likelihood that the translated version is equivalent to the original version of the item. The item is first translated into the target language and then independently translated back into the original language; the two versions are compared to identify any points of disagreement. However, as Stansfield (2003) points out, inconsistency can
be introduced either in the original translation process or in the back translation process. Also, if the initial translator knows that back translation is employed, there may be a tendency to use literal language rather than more natural expression in the translation. To create equivalent translations, being bilingual is not a sufficient qualification; professional translators should be employed, preferably with knowledge of basic principles of item writing. With or without back translation, successive review and revision of items is recommended (Stansfield, 2003). Back translation has been used successfully in developing an equivalent Spanish version of a Minnesota state assessment (Liu, Anderson, & Swierzbin, 1999).

Translated test items may be used to create a separate test booklet, or bilingual booklets may be created with the original and translated items (e.g., in English and in Spanish) side-by-side. In a state assessment in 1997, the Massachusetts Department of Education used Spanish language booklets as well as side-by-side bilingual booklets with Spanish and English on facing pages. Interviews with 997 students indicated that students who received bilingual test booklets relied mostly on the Spanish version of the items, but in some cases they also read the English version (Stansfield, 2003).

In a Minnesota State reading assessment, most students who used a bilingual version of the test booklet relied on only a single language; however, some students referred to both versions. Students did not report being distracted or confused by the format (Liu, Anderson, & Swierzbin, 1999). The use of the dual language test booklets showed a positive effect on test scores that was not statistically significant (Anderson, Liu, Swierzbin, Thurlow, & Bielinski, 2000).

In another study students were interviewed after taking an eighth grade mathematics test using bilingual test booklets. Students who were native Spanish speakers with more than three years’ instruction in the U.S. focused mostly on the English pages; some used the Spanish pages to check their understanding or to find the meaning of specific words or phrases. Students with fewer than three years’ instruction in the U.S. focused on the Spanish pages most of the time. Of the 181 students using the bilingual booklet, 85% indicated that the bilingual format was useful or very useful in comparison with an English-only test booklet (Garcia et al., 2000, as reported in Stansfield, 2003).
A meta-analysis by Francis et al. (2006) found that the two studies of dual language accommodations showed opposite effects. They suggest that the appropriateness of this accommodation may depend on the level of literacy and/or instruction in the students’ native language. The bilingual test booklet may be a viable option to address construct-irrelevant language challenges, especially in U.S. schools where most bilingual students share the same native language, such as Spanish. Bilingual booklets were selected as an accommodation for the current study.

**Accommodation by Linguistic Modification**

A number of studies have sought to reduce the language barrier for ELLs by modifying the vocabulary and linguistic structures of test items in content areas such as mathematics and science, but the results have been mixed. Minor changes in the language of mathematics word problems can affect student performance (Cummins, Kintsch, Reusser, & Weimer, 1988; DeCorte et al., 1985; Hudson, 1983, Riley, Greeno, & Heller, 1983). Mathematics test performance of some students has been affected by differences in the syntactic complexity of the language of word problems (Larsen, Parker, & Trenholme, 1978; Wheeler & McNutt, 1983). Abedi, Lord, and Plummer (1997) found that the performance of eighth grade ELL students on NAEP mathematics test items was lower than that of non-ELL students and that the difference was significantly greater on items with more words, regardless of the difficulty of item content.

A study by Abedi and Lord (2001) compared the performance of eighth grade students on NAEP mathematics items with parallel items that were modified to reduce the complexity of sentence structures and to replace potentially unfamiliar vocabulary with words likely to be more familiar to the students. The mathematics tasks were not changed, nor were mathematical terms in the items. In this study 1,031 students in classrooms were randomly assigned original versions and modified versions of the items. Test results showed small but significant differences in the scores of students in low level and average mathematics classes; students scored higher on the modified items. ELLs and low-performing students benefited the most from language modification. The linguistic features that appeared to contribute to item difficulty included low-frequency vocabulary and passive voice verb constructions (see Abedi, Lord, & Plummer, 1997, for discussion of linguistic features).
In another eighth grade study with 1,394 students in schools with high enrollments of Spanish speakers, students generally scored higher on shorter problem statements. Modification of the language of items contributed to improved performance on 49% of the items (Abedi, Hofstetter, & Lord, 1998). In a third study (Abedi, Lord, Hofstetter, & Baker, 2000) comparing various accommodations with eighth grade mathematics items, the modified English version narrowed the score gap between ELLs and other students.

Other studies employing linguistic modification of items have not found significant differences in student scores (Brown, 1999; Lotherington-Woloszyn, 1993; Rivera & Stansfield, 2001), possibly due in part to small sample size. Rivera and Stansfield (2001) found that linguistic simplification did not affect the scores of English-proficient students. This indicates that comparisons could be made using scores on accommodated and non-accommodated versions.

In a meta-analysis of several studies of linguistic modification, Francis et al. (2006) found that the overall average effect size for this accommodation was not significant. However, they note that some of the randomized studies using linguistic modification employed small sample sizes of ELLs. The effectiveness of linguistic modification as an accommodation depends on the nature of changes that have been made in the language of the items. Since different studies may make use of different modifications, it may not be informative to generalize across all modification studies. Because of its potential effectiveness and validity, this accommodation was employed in the current study.

**Accommodation Providing Extra Time**

The provision of additional time on a test is a popular accommodation for ELLs and students with disabilities. It is appropriate for students who may take longer to read, comprehend, and respond to test items. The empirical evidence for the effectiveness of this accommodation is mixed. Extended time may benefit students with learning disabilities under certain conditions, but it can also help students without learning disabilities (Jensen, 1997). Other studies indicate that some students with learning disabilities may not benefit differentially from extended time; some students may actually be advantaged if other students are not allowed
extra time (Fuchs & Fuchs, 1999; Fuchs et al., 2000; Elliott S. et al, 2001). A review of the literature concluded that the accommodation of extended time improved the performance of students with disabilities more than it improved the performance of students without disabilities (Sireci, Li, & Scarpati, 2003). This accommodation requires a closer look, particularly for the heterogeneous group of students classified as learning disabled, with respect to its effectiveness as well as validity. Some students apparently can benefit from extended time, even though group averages may not show significant differences (Stretch & Osborne, 2005; Cawthon, 2006).

For ELLs studies show that the provision of extra time can result in higher scores. However, this accommodation may also produce higher scores for non-ELLs, thus establishing its effectiveness but bringing its validity into question (Hafner, 2001; Kopriva, 2000; Abedi, Hofstetter, & Lord, 1998; Abedi, Kim-Boscardin, & Larson, 2000).

Another study indicates that extra time may not always be an effective accommodation. A study of eleventh grade ELLs using three accommodations found that student scores were lowest with the provision of extra time (Miller, Okum, Sinai, & Miller, 1999). A meta-analysis by Francis et al. (2006) found that extra time had a positive effect, but the effect size was not statistically different from zero.

In some studies, and in practice, students are provided extra time on assessments in conjunction with other forms of accommodation, and it becomes difficult to determine which effects, if any, are due to the provision of extra time. Access to a dictionary or glossary may not be effective alone, for example, but may produce higher scores if students are given extra time to consult the dictionary. Further study is needed to determine effectiveness and validity of this accommodation for ELLs, students with disabilities, and ELL students with disabilities.

Most articles on inclusion in large-scale assessment and accountability programs discuss ELLs and SWDs separately. Thurlow and Liu (2001a) argue that if ELL students with disabilities are left out, they are, in effect, deprived of the benefits of standards-based reform. The lack of attention to this group in the research literature has been paralleled by the absence of participation and performance data for these students in national, state, and district large-scale assessment data bases (Thurlow & Liu, 2001b). However, in recent studies these students have
begun to receive increased attention. The present study addresses a gap in the research, investigating the validity of assessments for ELL students with disabilities as well as ELLs and SWDs and focusing on the effectiveness and validity of assessment accommodations for this group of students.

**Methodology**

**Research Hypotheses/Questions**

While the immediate goal of this project is to help states in the assessment of English proficiency for ELL students with disabilities, the outcome of this study may inform future national and state assessments on determining appropriate accommodations for English language learners and students with disabilities in general and for ELL students with disabilities in particular. This study has the unique opportunity of involving a number of states, which gives strong generalization power to the results.

Issues concerning the use of accommodation for ELL students in the assessment of English language proficiency are more complex than accommodations in content-based areas such as mathematics and science. In content-based areas, the construct being assessed is not English language proficiency; therefore ELL students can be provided with accommodations that help with the language barrier. However, in English language proficiency assessments, language is the construct being measured; consequently, any accommodation that helps students with the English language content (such as use of glossary or simplified English) may affect the measurement of the construct, thereby invalidating the results.

Studies have been conducted recently to address the issues concerning the validity of some of the accommodations for both SWD and ELL students (Abedi, Courtney, & Leon, 2003; Abedi, Lord, Boscardian, and Miyoshi, 2003; Abedi, Lord, Hofstetter & Baker, 2000; Elliot, Braden, & White, 2001; Olson & Goldstein, 1997; Thurlow, Elliott, & Ysseldyke, 1998; Walz, Thompson, Thurlow, & Spicuzza, 2000). However, based on our literature search, to date there are no published experimental studies to examine validity of accommodations used for ELL students with disabilities. Thus, while participating states in the consortium for English language proficiency test development feel obligated to use accommodations for some of the ELL students
with disabilities who need the accommodations the most, they are uncertain on how to use the accommodated results without knowing much about the nature and validity of those accommodations. Unfortunately, many states have not had enough support for research on the key issues concerning assessment of English language proficiency.

There are several major issues that needed to be examined carefully before using accommodations in the assessments of SWD, ELL, and ELLWD students. Among the most important issues are the following:

1. How effective is the accommodation in increasing the performance of SWD, ELL and ELLWD students?
2. Are the accommodated assessments as valid as the non-accommodated assessments?
3. Do accommodations work the same for all SWD, ELL and ELLWD students with different background?
4. How feasible are the accommodations that are found to be effective and valid?

The focus of this study was on the validity of accommodations in English proficiency and mathematics assessments for these students. Students’ performance is defined as scores from ELDA and mathematics assessments. A number of student background variables were included in the study.

Population and Sample

The population for this study consisted of students in elementary, middle and high schools with the largest concentration of ELL students. The sample included students in grades 4 and 5 from elementary schools, grades 7 and 8 in the middle schools and grades 11 and 12 in high schools. The design of this study called for each of the participating states to provide access to two elementary, two middle and two high schools with each school providing access to four classes. In elementary schools each school was to provide access to two classes in grade 4 and two classes in grade 5. In middle schools each school was to provide access to two grade 7 and two grade 8 classes, and in each high school two grade 11 and two grade 12 classes were to be tested. However, some of the participating states were not able to deliver that level of
participation due to technical and logistical difficulties. Therefore, the issue of sample size—particularly with ELL students with disabilities became a major concern in this study.

At the proposal submission phase, 14 states indicated that they would participate in the study if the proposal is accepted. However, after the approval of the proposal many of the volunteered states decided to not participate mainly due to too much testing demand and too much assessment burden on students. They were concerned about losing instructional time to testing and consequently not being able to comply with the NCLB requirements. Some of the states that withdrew from the study informed the project team of their intention at the very beginning of the project. Nonetheless, a significant number of them stayed in the project and continued with their intention to participate until a few days prior to actual testing. These cases consumed a substantial amount of project time and resources.

Table 1
Total number of students tested in both ELDA and Math by states

<table>
<thead>
<tr>
<th></th>
<th>ELDA Test</th>
<th></th>
<th>Math Test</th>
<th></th>
<th>Total Students</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Pct Tot</td>
<td>N</td>
<td>Pct Tot</td>
<td>N</td>
<td>Pct Tot</td>
</tr>
<tr>
<td>California (LAUSD)</td>
<td>837</td>
<td>58.0</td>
<td>759</td>
<td>61.7</td>
<td>1,596</td>
<td>59.7</td>
</tr>
<tr>
<td>Kentucky</td>
<td>163</td>
<td>11.3</td>
<td>120</td>
<td>9.7</td>
<td>283</td>
<td>10.6</td>
</tr>
<tr>
<td>New Mexico</td>
<td>33</td>
<td>2.3</td>
<td>35</td>
<td>2.8</td>
<td>68</td>
<td>2.5</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>279</td>
<td>19.3</td>
<td>160</td>
<td>13.0</td>
<td>439</td>
<td>16.4</td>
</tr>
<tr>
<td>Oregon</td>
<td>40</td>
<td>2.8</td>
<td>56</td>
<td>4.5</td>
<td>96</td>
<td>3.6</td>
</tr>
<tr>
<td>Virginia</td>
<td>89</td>
<td>6.2</td>
<td>101</td>
<td>8.2</td>
<td>190</td>
<td>7.1</td>
</tr>
<tr>
<td>Missing</td>
<td>2</td>
<td>0.1</td>
<td>1</td>
<td>0.1</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Total</td>
<td>1,443</td>
<td>100.0</td>
<td>1,232</td>
<td>100.0</td>
<td>2,675</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 1 shows the number of students from each of the six participating states in this project. As data in Table 1 show six states took part in this study. The participating states are: California (Los Angeles Unified School District), Kentucky, New Mexico, Oklahoma, Oregon, and Virginia with a total of 2,675 students. However, the number and percent of students are very different across the participating states, ranging from 1,596 students in California to 68
students in New Mexico. Such large differences can be explained by the size of the population of states but mostly due to the willingness of individual school within state to cooperate. As indicated earlier, many states did not commit fully to this project, and even when they did they changed their level of participation due to the test burden on students. This wide range in the participation rate certainly affected our ability to generalize results to the participating states as well as to the nation. However, we welcomed the opportunity of selecting more subjects from California as it is the state with the largest number of ELL students, having a third of the ELL population in the nation (Kindler, 2002).

Table 2 presents distributions of the sampled students by their background characteristics. As indicated earlier, a total number of 2,675 students were tested in the following six states: (1) California (Los Angeles Unified School District), (2) Kentucky, (3) New Mexico, (4) Oklahoma, (5) Oregon and (6) Virginia. As the data in Table 2 show, of this number, 555 (20.8%) students were in grade four, 521 (19.5%) in grade five, 458 (17.1%) in grade seven, 711 (26.6%) in grade eight, 232 (8.7%) in grade eleven and 196 (7.3%) in grade 12. Over forty seven percent of the students were female and 52.7% were male. Most of the students in the sample were recipients of free-reduced price lunch program (80.6%). Over 67% of the students were Latinos, 24.1% were Caucasian and 8.4% were African Americans. Of the total sample, 401 or 15% had some form of disabilities. Due to our emphasis on the ELL population representation, over 47% of the students in the sample were English language learners.
Table 2
Total number of students tested in both ELDA and Math by student background variables

<table>
<thead>
<tr>
<th></th>
<th>ELDA Test</th>
<th>Math Test</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>% of Tot</td>
<td>N</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>639</td>
<td>44.3</td>
<td>546</td>
</tr>
<tr>
<td>Male</td>
<td>652</td>
<td>45.2</td>
<td>670</td>
</tr>
<tr>
<td>Missing</td>
<td>152</td>
<td>10.5</td>
<td>16</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latino</td>
<td>774</td>
<td>53.6</td>
<td>851</td>
</tr>
<tr>
<td>Caucasian</td>
<td>431</td>
<td>29.9</td>
<td>205</td>
</tr>
<tr>
<td>African America</td>
<td>121</td>
<td>8.4</td>
<td>99</td>
</tr>
<tr>
<td>Asian</td>
<td>35</td>
<td>2.4</td>
<td>25</td>
</tr>
<tr>
<td>Other</td>
<td>62</td>
<td>4.3</td>
<td>31</td>
</tr>
<tr>
<td>Missing</td>
<td>20</td>
<td>1.4</td>
<td>21</td>
</tr>
<tr>
<td><strong>Free/Reduced Lunch</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1047</td>
<td>72.6</td>
<td>982</td>
</tr>
<tr>
<td>No</td>
<td>263</td>
<td>18.2</td>
<td>225</td>
</tr>
<tr>
<td>Missing</td>
<td>133</td>
<td>9.2</td>
<td>25</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4&lt;sup&gt;th&lt;/sup&gt;</td>
<td>555</td>
<td>20.8</td>
<td>0</td>
</tr>
<tr>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0</td>
<td>0.0</td>
<td>521</td>
</tr>
<tr>
<td>7&lt;sup&gt;th&lt;/sup&gt;</td>
<td>458</td>
<td>17.1</td>
<td>0</td>
</tr>
<tr>
<td>8&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0</td>
<td>0.0</td>
<td>711</td>
</tr>
<tr>
<td>11&lt;sup&gt;th&lt;/sup&gt;</td>
<td>232</td>
<td>8.7</td>
<td>0</td>
</tr>
<tr>
<td>12&lt;sup&gt;th&lt;/sup&gt;</td>
<td>196</td>
<td>7.3</td>
<td>0</td>
</tr>
<tr>
<td><strong>Disability-status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not disabled</td>
<td>1218</td>
<td>84.4</td>
<td>1056</td>
</tr>
<tr>
<td>Has Some Disability</td>
<td>225</td>
<td>15.6</td>
<td>176</td>
</tr>
<tr>
<td><strong>ELL status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-ELL</td>
<td>625</td>
<td>43.3</td>
<td>551</td>
</tr>
<tr>
<td>ELL</td>
<td>520</td>
<td>36.0</td>
<td>541</td>
</tr>
<tr>
<td>Missing</td>
<td>298</td>
<td>20.7</td>
<td>140</td>
</tr>
</tbody>
</table>

Students were tested within intact classrooms. Unlike the percentages of students with disabilities that vary slightly across the nation (ranging from 8% to 14%), percentages of ELL students differ greatly across the nation. For example, over 27% of the public school students in California are English language learner, while this percentage drops to less than one percent for
many other states (Kindler, 2002). Due to such substantial levels of variation and because the ELL student population was small in some of the participating states, we requested access to schools with the largest number of ELL students in the state.

While we have no information to suggest that the sample for this study represents different population characteristics than the national sample, we do not intend to generalize the findings of this study beyond what the current sample permits.

**Instruments**

The design of the study called for the re-administration of ELDA and the state mathematics test. With regard to the state mathematics test, there were concerns over re-administering the state-developed mathematics tests for three main reasons: 1) the test publishers might not feel comfortable using the tests for research purposes; 2) due to the effect of testing, students who are tested twice may show improved performance over those tested for the first time; 3) and most importantly, it is an additional burden to the participating states. In order to deal with these issues an alternative plan was devised and implemented. Rather than using the state mathematics test for each participating state, a mathematics test for each of the grade clusters was created using NAEP released mathematics items.

In order to evaluate the content coverage of the mathematics test, Advance Research and Data Analyses Center (ARDAC) – the organization responsible for implementing the study-conducted an alignment study to make certain the content of the mathematics test was appropriate for the participating states. In doing so ARDAC contacted states in collaboration with the Oklahoma State Department of Education (OSDE). Of the eleven states, seven (Georgia, Iowa, Kentucky, Louisiana, Nevada, New Jersey, and Oklahoma) chose to participate in the alignment study. For the alignment study, ARDAC selected 59 NAEP-released mathematics test items for grade 4 and 67 NAEP-released mathematics test items for grade 8. All mathematics test items were presented to each participating state in addition to a test item information summary sheet which included details regarding the question type, percentage of students who responded correctly, mathematics content area and mathematics ability. Additionally, a column was assigned for the rating scale where each NAEP-released mathematics test item was given an alignment rating against state content standards for each participating state. The rating scale was
from 1 to 5, with 1 indicating no alignment and 5 indicating maximum alignment.

Research staff at ARDAC conducted the alignment study with each of the seven participating states and individually recorded the alignment ratings. Once all alignment ratings were received, ARDAC calculated the average of all seven ratings for each NAEP-released mathematics test item. Upon calculating the averages, the top highly aligned 30 mathematics items were selected to be incorporated in the math test to be administered for the study.

The ELDA accommodations are the accommodations that are specified in a student's IEP or 504 plans. This information is provided by the school coordinator of the participating schools. Any accommodations for students with specific disabilities must be specified before they take the assessment. Only those accommodations specified in a student's IEP or 504 Plan may be provided during the assessment.

**Student background questionnaire, teacher questionnaire, and school questionnaire**

The student background questionnaire (Appendix A), teacher questionnaire (Appendix B), and school questionnaire (Appendix C) were developed for this study to provide demographic and self-reported proficiency in English and the student’s native language to complement test results. All questionnaires were field tested on several classes at Santa Monica High School in California. These classes completed the questionnaires and provided feedback on them. Students taking the test completed the questionnaire and provided information such as the student’s perception of their skill in speaking, writing and understanding English; the ZIP code (type of community) in which they lived; the number of times they had switched schools due to moving, their birthplace and time spent living in the U.S., the language spoken at home and how often it was used, and the educational level of accomplishment they expected they would attain.

Teachers whose classes were tested completed the teacher questionnaire and provided requested information such as the teacher’s perception of students’ English writing, reading and speaking proficiency; teacher education and certification level; tenure at the school; teaching experience; and ethnicity. In addition, teachers were requested to provide more information on the specific accommodations provided to English language learner students and students with disabilities and on the use of these accommodations in instruction, testing and assessments.
Principals whose schools were tested completed the school questionnaire and provided requested information such as classification as LEP (limited English proficient), FEP (fluent English proficient) and IFE (initially fluent English proficient); free/reduced-lunch eligible; and proportion of students with disabilities within their school population; size of school library; and availability of tutoring, after school activities and computer labs.

Research Design

Among the four major issues in the use of accommodations (effectiveness, validity, differential impact and feasibility), this study focused on the effectiveness and validity of accommodations. For the Oklahoma State Department of Education (OSDE) and collaborating states effectiveness and validity issues are the main concern in the use of accommodations for any subgroup of students (ELL, SWD, and ELLWD). To examine the effectiveness of accommodations performance of accommodated ELL, SWD, and ELLWD students were compared with the performance of non-accommodated ELL, SWD, and ELLWD students. Validity of accommodations for these students were tested using two different approaches: 1) comparing the performance of accommodated and non-accommodated students who are not in the ELL or SWD category; and 2) comparing the validity of accommodated assessments with non-accommodated assessments using a criterion-related validation approach. Results from these two approaches can be used for cross-validation purposes.

The design for this study was quasi-experimental, since intact classrooms were used; individual students were nested within classrooms and classrooms were nested within districts and states. Thus, in addition to variability due to individual student differences, there was variability due to classroom, district and state differences. Accordingly, the design for this study was hierarchical. However, since students were assigned randomly to different accommodation conditions within the classroom, the higher order variables (classroom, district and state) may have limited effect on the research outcomes.

Random assignment of students to different accommodation conditions within the classroom controls for many of the sources that may threaten the internal validity of the design (see McMillan & Schumacher, 1997, pp183-190). Among the most important sources of threat
to internal validity is “selection,” which in this case refers to the initial differences between the accommodated and non-accommodated groups. The most important sources contributing to initial differences between the two groups could be teacher, school and district effects. By using random assignment of students to experimental conditions within classrooms, the possibility of such initial differences was greatly reduced. Similarly, random assignment of students to accommodation conditions within classrooms decreases the threat level to internal validity of the design due to “history”, ”attrition” and “instrumentation” (see McMillan & Schumacher, 1997, pp183-190).

Below are the null hypotheses for examining the validity of accommodations, one for performance comparison, and one for the criterion related validity approach. The term “accommodated condition” used in this proposal refers to assessments under the state-approved accommodations for ELL students with disabilities.

H₀₁: Non-ELL/non-SWD students will perform the same under accommodated and non-accommodated conditions.

H₀₂: Structural relationships between the ELDA and mathematics assessment outcome measures and the criterion variables are the same for the accommodated and non-accommodated assessments.

This study may be replicated by states that are not participating in this study. The experimental design of this study (the quasi-experimental with random assignment of accommodations to different groups) can clearly be applied to similar situations in other places with similar content assessments.

**Procedure**

As indicated earlier, the purpose of this study was to examine the effectiveness of the state-approved accommodations in reducing the performance gap between special needs student populations (ELL, SWD and ELLWD) and their non-ELL/non-SWD peers. The study also examined the validity of the accommodated assessments. Our initial plan was to examine the effectiveness of accommodation and the validity of those accommodated assessments for ELL
students with disabilities only. However, to assist the participating states with their accommodated assessments in a broader sense, we included English language learners and students with disabilities as separate groups in the sample. Further, the initial plan was to examine accommodation issues in the assessment of English language proficiency (ELDA) but we also expanded the scope of this study to include mathematics because of the potential benefit for states, since mathematics is a major content area on which accommodation research may benefit states.

We obtained detailed information needed to assign appropriate accommodations to students through a Class Roster (see Appendix D). The Class Roster form was faxed or mailed to teachers as soon as school approved participation. Teachers were asked to complete and fax or mail the Class Roster to ARDAC a week before the test administration. The front page of Class Roster asked for teacher and school name, grade level, and name of the district and state. The Roster then asked for background information and assessment results for each student in the class, including gender, ethnicity, student’s disability, ELL status, and free/reduced price lunch program participation. The assessment results include state test scores on mathematics and reading and teacher’s rating on mathematics and reading. As indicated earlier, these data were used as external criteria for examining the validity of accommodated assessments. More importantly, the Class Roster asked the teacher to provide information on the type of accommodations that students with disabilities should receive based on their IEP or 504 plans. The Class Roster provided instructions for the teacher about the information that was requested in the Roster (e.g., categories for ethnicity, disability status and ELL status, see Appendix D).

As discussed earlier, state-approved accommodations that were identified in students IEP or 504 plans were provided to students with disabilities to the extent possible. Appendix E provides a list of state approved accommodations some of which were used in this study. For ELL students a limited number of accommodations were provided. ELL students received accommodations only in the mathematics test. Since most of these accommodations were language-related, they were not appropriate for the English language proficiency assessment. These accommodations included a linguistically modified version of the mathematics test, large print, extended time and bilingual booklet of the mathematics test. To estimate the reliability of the mathematics test used in this study and to examine the stability of scores across different
forms, the mathematics test was presented in two forms (Forms A and B). These forms were considered as two parallel forms. They both included exactly the same items (identical). The only difference between the two forms was the order of items. The order of items in Form B was the opposite of Form A.

**Contacting the schools**

Originally, it was the states’ responsibility to contact schools with a high percentage of ELL students (high density ELL schools). When some states proved unable to do so, ARDAC researchers contacted High Density ELL schools and asked them to participate in this study. Schools agreeing to participate appointed a school contact person to correspond with ARDAC, select classes, send and receive documents and test materials, and select test administrators from the school sites. School contact persons were paid $250 to compensate for their time. While originally ARDAC staff were to administer testing, it was more feasible for school personnel to function in this role. Through extensive correspondence and a teleconference training session, school contact persons and test administrators were assisted in selecting classes, providing the required information, preparing testing conditions, distributing test materials, and administering testing. ARDAC maintained contact with these schools through the testing process. At the end of the process, the schools received compensation of $250 for participation.

**Preparing for testing**

To ensure exact implementation of the complex study design, ARDAC staff and school site personnel engaged in detailed preparations for the test administration. Each school contact person was given a parent assent letter (Appendix F) and a Class Roster. The parent assent letters were to be distributed to the parents of students in the selected classes to provide them an opportunity to request that their child be exempted from participation if they choose to do so. Children of the parents expressing concerns about their children participation in the study were excluded from the study. Parent assent letters were also translated into Spanish and Russian, and teachers were asked to distribute this letter at least a week before testing so that booklets would not be prepared and sent to students who would not be participating.

The Class Roster (Appendix D), when filled out, would list students’ names, gender,
ethnicity, LEP status, language spoken, disability status, receipt of free school lunch, state mathematics and reading scores, and teachers’ rating of students’ reading and mathematics abilities. A sheet listing students who required state-approved accommodations according to their IEP or 504 Plan and the accommodations students should receive was attached to the Class Roster. Most students who require test accommodations have an IEP or 504 and detailed instructions for their accommodation. Teachers were asked to list any and all accommodations their students receive on the Class Roster, including their type of disability. Based on all of the information mentioned above, ARDAC research staff prepared test booklets with the name of students written on the booklet prior to test administration. For example, it was clear who receives which booklet with which accommodations and under which condition.

We compared the requested accommodations with our own list of approved accommodations for this study to determine which of the accommodations would provide valid outcomes. This information was also used to prepare the test administrator and the school’s staff for the anticipated extra facilities, scheduling flexibility, and assistance needed for the day of the test. We requested the most detailed information possible about these students’ needs and thoroughly followed up with instructors, school contact persons, and test administrators for more information to ensure that our preparation for accommodations was complete and that the students would be able to take the test under the conditions they required.

For either test being administered SWD students were assigned a “twin,” a student selected from the regular student group in the class who had similar state test scores and teacher ratings in the subject being tested for that class. To test the effectiveness of accommodations received, one half of the students (randomly selected) requiring accommodations in each class would receive their accommodations while the other half would not. The twin for each student would receive the same testing conditions as the SWD student. To test the validity of accommodation, performance of twins (non-ELL/non-SWD who received accommodations) were compared with a similar group of non-ELL/non-SWD who did not receive any accommodations. School contact persons and test administrators were informed in advance of what accommodations they would need to provide during testing.

For grades taking the mathematics test, accommodated students listed as LEP were
assigned a twin who was not SWD but who spoke the same language and had similar state test scores, teacher rating, and LEP rating. The distribution of the different formats of the mathematics test was also determined according to the Class Roster. Spanish-speaking students listed as ELL who were not SWD were randomly assigned either a regular English language mathematics booklet, or a bilingual side-by-side English-Spanish mathematics booklet. Among the remaining students not receiving accommodations or a bilingual test booklet, one half was randomly selected to receive a simplified version of the mathematics test. After these booklet assignments had been determined, ARDAC staff labeled each booklet with the student’s names, grades, teachers, schools, test forms received, and any accommodations that student would receive.

*Training test administrators*

In view of the difficulties associated with ARDAC staff administering tests at all school sites, test administrators were selected from the schools’ staff and requested to participate in teleconference training. The training took place at least several days before testing once the Class Rosters had been prepared and the accommodations for SWD and LEP students assigned. Using a summary of the accommodations each class would receive, the timing instructions for each test, and a list of packed materials, ARDAC staff discussed with the test administrator the testing process from receipt of the booklet shipment to return of completed tests. ARDAC familiarized test administrators with the processes for delivering the most common accommodations and discussed with them their plan for testing, including their room reservations, test scheduling, and use of assistants for testing according to their accommodated students’ particular needs. Processes for maintaining test security were heavily stressed, and administrators were instructed to contact ARDAC upon testing completion for instructions in shipping their completed materials. Test booklets labeled for individual students were sent to schools, and ARDAC contacted schools to get confirmation that the materials had been received.

*Test administration*

The procedure for administering the ELDA test was taken from the ELDA’s own directions for administration and provided to test administrators at the schools. For administering
the mathematics test, ARDAC developed a set of instructions for schools to use when administering the mathematics test and included it in the shipment of test materials for schools being tested in mathematics. In addition to reading and understanding the directions for testing, test administrators were expected to review the *Test Administrator Training Agenda* (Appendix G) before participating in the teleconference training prior to their testing date. The *Agenda* included a short summary of the tests with their timing constraints and a list of facilities, assistance and materials, including the shipment of booklets and questionnaires from research office that schools would need on the day of the test. Additionally, a summary of the accommodations needed for each class was included in the agenda document, intended to help administrators prepare in advance any special testing conditions, facilities or assistance they would need. At some schools, ARDAC researchers participated in test administration, either through administering the test or supporting the school’s test administrator.

**Test scoring**

When schools completed ELDA testing and student background questionnaires, the ELDA booklets were sent to Measurement Incorporated (MI) and the questionnaires were sent to ARDAC. MI was responsible for scoring completed test booklets and returning data to ARDAC to be merged with the student, teacher and school background information. When schools completed mathematics testing and background questionnaires, the mathematics booklets and the questionnaires were sent to ARDAC. We matched the names on the labels of the mathematics books with the labels on the student background questionnaires. ARDAC processed and coded the mathematics booklets and background questionnaires based on the pre-established coding systems.

The student background questionnaires for both ELDA and mathematics classes were matched with the student information from the Class Roster before coding. Coding of the questionnaire was completed using a master sheet that assigned a number for each column on the Class Roster and continued on to provide sections to record the responses to the questionnaire items. Numerical codes were assigned for the state, school, class, student, ethnicity, absence, receipt of parent assent, disability type, LEP status, receipt of free lunch, state mathematics score, state reading score, teacher mathematics rating, and teacher reading rating. The
questionnaire items were also assigned a number, and the students’ responses were recorded according to a code as well. As with the student questionnaire, the teacher questionnaire items were numbered for coding, and the teachers’ responses were recorded using numerical code. Additionally, the responses to the teacher questionnaire items regarding students’ disabilities were recorded using numbers that corresponded to the disabilities and the accommodations provided.

The coding for ELDA tests was done by MI, while the coding for the mathematics test was performed by ARDAC research staff. For classes taking the mathematics test, additional coding columns were added to record the students’ responses to the items. A number was assigned for the test the student took, the version of their test, the question number, each answer choice for multiple choice items, and the response students provided to open-ended questions.

Creation of the Linguistically Modified Version of Mathematics Items

Mathematics tests were created for grades 5 and 8, with 30 items in each test. Items were reviewed for sources of potential linguistic difficulty. Several of the items were found to contain linguistic features that previous research has suggested may constitute difficulty or may take more time and/or cognitive resources to process. These features included subordinate clauses, unfamiliar vocabulary (low-frequency words, typically long words), passive voice constructions, modal verbs, and participial modifiers. Selected items were rewritten, with the objective of reducing the incidence of these features as well as reducing sentence length and item length. The original mathematics concepts were retained in the modified items as were the mathematics tasks and computations required by the original items.

For grade 5, modified versions were constructed for 15 of the 30 items in the test. Nine of the 15 items were rewritten to contain fewer clauses per sentence (i.e., fewer subordinate clauses). Eight of the modified items used simpler vocabulary. Five items were rewritten to remove passive voice constructions, three to remove modals, and five to remove participial modifiers. For the grade 5 assessment, all 15 modified items had an average sentence length that was shorter than in the original items, and 13 of 15 modified items contained fewer total words than the corresponding original items.
For grade 8 of the 15 modified items, eight were rewritten to contain fewer clauses per sentence. Nine items were rewritten using simpler vocabulary. Four items were rewritten to remove passive voice constructions, two to remove modals, and six to remove participial modifiers. In grade 8 modified items, the average sentence length was shorter than in the corresponding original items, and 10 of the 15 modified items contained fewer total words than their original counterparts.

For each grade, a test booklet was created containing the 30 original mathematics items. Another test booklet was created including the modified versions of 15 items in addition to the remaining 15 items with no modifications.

**Item Ordering and Large Print Version**

Two forms of the mathematics test booklet, Form A and Form B were created. The two forms differed only in the order in which the items were presented. The rationale for two forms was based on the observation that students sometimes fail to complete a test, with items at the end more likely to be left unanswered or unattempted. In this circumstance, an item that was not challenging could be left unanswered, resulting in a zero score on that item. This fact could distort the validity of the resulting p-value as an indicator of item difficulty. Using two test forms with different item orders would help to minimize the possible influence of this factor.

A version of the test booklet with large print was also created because this accommodation was one of the most widely-used accommodations by states and because the results of previous research on the effectiveness and validity of large-print test format were mixed.

**Bilingual Test Booklet**

A bilingual version of the mathematics test booklet was created, with individual test items being translated from English to Spanish. Then, as a check on accuracy, other consultants translated the Spanish versions back to English, and the results were compared with the original English versions. Spanish items were edited to ensure comparability with English items with respect to mathematics content and language difficulty. In the bilingual test booklet, English and
Spanish versions were presented side-by-side.

Two translation agencies (Better Communication and Ángeles de la Fuente Spanish) were asked to translate the grade 5 and 8 math tests into Spanish for Spanish speaking students, then back-translated the Spanish version into English (for accuracy). We contacted several schools to find 5th and 8th grade math teachers who are fluent in both English and Spanish to compare the two English versions. Teacher compared the two English versions for each grade to make sure that there are not major differences between the two versions that could impact on the math construct being measured. In addition we provided a rating sheet and asked the teachers to make suggestions and comments for improvement of the version in case of inconsistencies between the two forms. We also asked two experts to review the translations. They reviewed the Spanish versions and offered their comments and advise and the translation agency incorporated their comments into the Spanish versions.

Accommodations: Effectiveness and Validity Issues

The efficacy of accommodations in the assessment of the special needs student population (SWD, ELL and ELLWD) is judged based on several factors, among the most important of which are effectiveness of the accommodations in reducing the performance gap between the special needs student population and their peers and the validity of the accommodated assessments. Accommodations are judged to be useful if they help to make assessment more accessible to the targeted population (effectiveness) thus reducing the performance gap between those in need of accommodations and those who do not need accommodations. However, this reduction in the performance gap is acceptable only if accommodations do not provide unfair advantage to the recipients (validity issues). In this study, we used special methodology to enable careful examination of these two important characteristic of the accommodations. To examine the effectiveness of accommodations, we had to provide accommodations to some of the students who are in clear need of receiving those accommodations and at the same time deprive others who are also judged to be in need of the accommodations from receiving accommodations. Furthermore, to examine the validity of accommodated assessment, we had to provide accommodations to those who do not need them so we could see if providing accommodations actually helped them in improving their
performance outcomes. Below, we will explain our approach in examining these two important aspects of accommodations.

The validity and effectiveness of accommodations were tested in two content areas, English proficiency and mathematics, and for three subgroups among the special needs student population (SWD, ELL, and ELLWD). For students with disabilities (SWD and ELLWD), we selected accommodations based on student IEP or 504 plans. The information about recommended accommodations was obtained from the Class Roster form which was completed by teachers prior to test administration (see Appendix D). Many different accommodations were suggested, but due to limitations in the number of accommodated students only a few of the most commonly suggested accommodations were provided.

In examining the effectiveness of accommodations and validity of accommodated assessments, we created four accommodation groups. Figure 1 shows these groups who are assessed under different accommodation conditions.

Figure 1.

Effectiveness and Validity of Accommodations

<table>
<thead>
<tr>
<th>Condition/ Accommodated</th>
<th>ELL/SWD/ELLWD</th>
<th>Non-ELL/non-SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accommodated</td>
<td>Group 1</td>
<td>Group 3</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>Group 2</td>
<td>Groups 4</td>
</tr>
</tbody>
</table>

Students in Group 1 are those who are in need of accommodations and received accommodations. Students in Group 2 are also in need of accommodations but were not provided accommodation in order that they serve as a control group to help us understand the impact of accommodations on the assessment of these students. A parental assent form (Appendix F) was sent to the parents of these students informing them of the necessity imposed by the research design for not providing accommodations to some of these students. Groups 3 and 4 do not need any accommodations as they do not have any disabilities or urgent need for language assistance. However, students in Group 3 received accommodations to determine any unintended effects due
to accommodations (any changes in the construct being measure). If there were changes in the construct being measured then by comparing the performance of students in Groups 3 and 4, such changes could be observed. We discussed this methodology in more detail below in the Validity and Effectiveness sections.

Examining the Validity of Accommodated Assessments

Two major hypotheses were tested to provide evidence on the validity of state-approved accommodations used in this study: 1) performance difference of non-ELL/non-SWD students under accommodated and non-accommodated conditions, and 2) comparing the criterion-related validity of the accommodated and non-accommodated assessments.

1. Performance Difference Approach

If accommodations do not alter the construct being measured, performance of non-SWD/non-ELL students (for whom accommodations are NOT intended) should remain the same under accommodated (Group 3) and non-accommodated (Group 4) conditions (Figure 1). Thus, a significant performance change under accommodations for the non-SWD/non-ELL group may suggest that accommodations are doing more than what they are supposed to do. To test this hypothesis, the mean scores of ELDA and mathematics of non-ELL/non-SWD were compared across the accommodation conditions (accommodated/ non-accommodated) using an analysis of covariance model in which students’ state reading score was used as the covariate. If the null hypothesis of no performance difference between the accommodated and non-accommodated students is not rejected, one can assume that the accommodation did not alter the constructs being measured. The null hypothesis is that the mean scores for ELDA and mathematics were the same for accommodated and non-accommodated students. That is:

\( H_0: \text{for validity of accommodation for ELDA:} \)

\[ M_{G3} (M_{ELDA}) = M_{G4} (M_{ELDA}) \]

\( H_0: \text{for validity of accommodation for mathematics:} \)

\[ M_{G3} (M_{Math}) = M_{G4} (M_{Math}) \]
It must be added, however, that due to the hierarchical nature of the design we also applied hierarchical liner model (HLM) model to the data of this study using teacher background variables (Appendix B) as the level 2 variables and school variable as level 3 variables.

The Results section will present evidence regarding these null hypotheses.

2. Criterion-Related Validation Approach

Comparing performance of non-ELL/non-SWD under accommodated and standard testing conditions may provide evidence on the validity of accommodations. However, examining the validity of accommodations using the criterion-related validity approach may produce even more convincing evidence for judging the validity of accommodations used in this study. We used both approaches in this study.

As indicated earlier, additional data were collected from students through the Class Roster (Appendix D). These data included student test scores from state assessments in mathematics and reading as well as teachers’ ranking of student performance in the two content areas. Students’ state assessment scores in mathematics and reading language arts as well as using some background variables such as self-reported proficiency in L1 and L2 and other academic background variables served as criteria to judge the validity of assessments and to compare structural relationship between accommodated and non-accommodated assessment for any possible trend differences between accommodated and non-accommodated assessment of special needs students and their peers. These variables were correlated with the ELDA and mathematics assessment scores in a latent-variable modeling framework. Using a multiple-group confirmatory factor analysis approach, structural relationships between the criterion variables and ELDA and mathematics test scores were compared across the accommodated and non-accommodated assessments. Several hypotheses of differences between the two groups were examined by testing the invariance of the relationship.

For example, Figure 2 presents the structural model for testing relationships between the four math item parcels and the math latent variable. This model was used once for accommodated and once for non-accommodated assessments using multiple group confirmatory factor analysis. As Figure 2 shows, math test items were grouped into four parcels. In this item parceling,
information on homogeneity of items within and between parcels were used in constructing parcels (for details on item parceling, see for example, Abedi, 1997). From the four item parcels a math latent variable was created and the following hypothesis of invariance was tested:

\[ H_0: \text{Factor loadings of item parcels with the math latent variables were the same across accommodated and non-accommodated assessments.} \]

A similar design can be applied to ELDA assessment. If the null hypothesis of invariance between the accommodated and non-accommodated assessments is NOT rejected, then one may assume that the structural relationship between assessment components in the accommodated and non-accommodated assessments remained the same which might be an indication of construct invariance between the two assessments.

**Examining the Effectiveness of Accommodations**

Effective accommodations should increase the performance of students who are the intended target of accommodations without changing the construct being measured. Under an effective accommodation the targeted students should perform better as compared to similar students under the standard testing condition with no accommodations offered. Therefore, comparing the performance of special needs students (ELL, SWD, ELLWD) under accommodations (Group 1) with a control group of similar students under the standard condition (Group 2, Figure 1) helped to examine the effectiveness of accommodations. If the accommodated group (Group 1) performed significantly higher than the control group under the standard testing condition (Group 2) in ELDA or mathematics, this may suggest that the accommodation was effective in increasing performance of this group of students. The null hypothesis is that the ELDA means for Group 1 (accommodated students) were the same as the means for Group 2 (non-accommodated students). That is:

\[ H_0: \text{for effectiveness of accommodation for ELDA:} \]

\[ M_{G1} (M_{ELDA}) = M_{G2} (M_{ELDA}) \]
H$_0$: for effectiveness of accommodation for mathematics:

$$M_{G1} (M_{Math}) = M_{G2} (M_{Math})$$

This null hypothesis must be rejected in favor of the following alternative hypothesis to suggest that accommodations used in this study were effective for ELL/SWD students:

$$M_{G1} (M_{ELDA}) > M_{G2} (M_{ELDA})$$

$$M_{G1} (M_{Math}) > M_{G2} (M_{Math})$$

To test this hypothesis, an analysis of covariance approach (ANCOVA) was used. Results of these analyses are presented below.

Results

As elaborated earlier in this report, the main purpose of this study was to examine effectiveness and validity of accommodations used in the assessment of English language learners with disabilities in the English language proficiency content area. We extended the scope of this project in two ways: adding the mathematics content area since it is one of major content areas that is included in the state and national assessments and accountability system (e.g. NCLB) and expanding the population to include both English language learners (ELL) with no apparent disabilities and non-ELL students with disabilities. Including more content areas and expanding the population to include broader categories of student gave us a better understanding of the nature of accommodations used for the special needs student population.

First, we will present an overall summary of students’ performance in ELDA and mathematics and then focus on the performance of subgroups of students in these content areas. For ELDA we used the raw scores provided by Measurement Incorporated (MI) and for the math, we transformed scores in Normal Curve Equivalence (NCE) scale with a mean of 50 and standard deviation of 21.06 (see Linn & Gronlund, 1995).

Table 3 presents the mean, standard deviation, and number of students taking ELDA reading and writing from the six participating states grades 4, 7, 9 and 11. As data in Table 3
show, there are discrepancies on the level of student performance from the different states. The mean performance across the states ranged from 29.29 (SD=12.22) to 38.32 (SD=17.87). Similarly, the mean writing scores ranged from 14.79 (SD=5.69) to 19.25 (SD=7.30). Again, for some states these means were computed based on a small sample size, and this could explain some of the variations.

Table 3
Mean, standard deviation and number of subjects tested in ELDA in all six states

<table>
<thead>
<tr>
<th>States</th>
<th>Reading</th>
<th></th>
<th></th>
<th>Writing</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>N</td>
<td>M</td>
<td>SD</td>
<td>N</td>
</tr>
<tr>
<td>California (LAUSD)</td>
<td>35.09</td>
<td>14.06</td>
<td>790</td>
<td>17.72</td>
<td>8.96</td>
<td>790</td>
</tr>
<tr>
<td>Kentucky</td>
<td>37.62</td>
<td>11.9</td>
<td>152</td>
<td>19.25</td>
<td>7.30</td>
<td>152</td>
</tr>
<tr>
<td>New Mexico</td>
<td>35.42</td>
<td>10.35</td>
<td>31</td>
<td>16.48</td>
<td>4.94</td>
<td>31</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>38.32</td>
<td>17.87</td>
<td>253</td>
<td>18.57</td>
<td>9.88</td>
<td>253</td>
</tr>
<tr>
<td>Oregon</td>
<td>29.29</td>
<td>12.22</td>
<td>38</td>
<td>14.79</td>
<td>5.69</td>
<td>38</td>
</tr>
<tr>
<td>Total</td>
<td>36.32</td>
<td>14.51</td>
<td>1346</td>
<td>18.28</td>
<td>8.72</td>
<td>1346</td>
</tr>
</tbody>
</table>

As discussed earlier in this report, for the mathematics test a set of released NAEP items was selected. Based on interviews with state assessment experts, those items aligned to the content standards of the participating states were used in this study. A mathematics test of 30 items (29 multiple-choice and one short-constructed response) was prepared for grade 5 students and a test of 30 items with four constructed-response items were prepared for students in grade 8. Table 4 presents mean, standard deviation, and number of students) in mathematics for grades 5 and 8 by states. Scores are reported in NCE scale with a mean of 50 and standard deviation of 20.06.
Table 4
Mean, standard deviation and number of subjects tested in math in all six states

<table>
<thead>
<tr>
<th>States</th>
<th>Grade 5</th>
<th></th>
<th>Grade 8</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>California (LAUSD)</td>
<td>47.76</td>
<td>19.16</td>
<td>392</td>
<td>40.11</td>
</tr>
<tr>
<td>Kentucky</td>
<td>58.52</td>
<td>22.78</td>
<td>87</td>
<td>53.75</td>
</tr>
<tr>
<td>New Mexico</td>
<td>57.69</td>
<td>21.64</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Oklahoma</td>
<td>58.51</td>
<td>19.13</td>
<td>65</td>
<td>67.39</td>
</tr>
<tr>
<td>Oregon</td>
<td>38.15</td>
<td>23.79</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>50.04</td>
<td>21.05</td>
<td>631</td>
<td>50.00</td>
</tr>
</tbody>
</table>

As data in Table 4 show, the mean math score is different across the participating states. The mean ranges from 38.15 (SD=23.79) to 58.52 (SD=22.78) with over 20 score points difference between states. Similarly, for grade 8 students there is a large performance difference between states. The mean NCE score ranges from 40.11 (SD=17.02) to 67.39 (SD=17.24). These differences can be due to large differences in sample size and different in the population characteristics between states.

In this study, in addition to the accommodations that were recommended for students with disabilities from their schools (based on IEP and 504 plans), we used three research-supported accommodations: 1) linguistic modification of assessment, 2) large print, and 3) bilingual version of the assessment (in mathematics). First, we will present the performance of students with disabilities and ELL students under these accommodations and then summarize the findings of accommodated assessments for these students under state-recommended accommodations.

To examine the effectiveness and validity of the three accommodations mentioned above for the participating states, the following five different forms of the mathematics tests were created for both grades 5 and 8 (see Appendix L for a link to all math test versions for grade 5 and Appendix M for all math tests for grade 8).

1. Original form without any changes in the content or format;
2. linguistically modified version in which the unnecessary linguistic complexity that prevents students’ (specially for ELL students) from understanding of the questions was
reduced;
3. large-print in which the questions were presented in larger print fonts;
4. linguistically modified form with the large fonts; and
5. bilingual booklet, a bilingual Spanish version of the mathematics test.

Two parallel forms of mathematics tests (Forms A & B, differing only in item order) were created for each of the five versions of the assessments. Since there were no statistically significant differences between students’ performance across the forms, we combined scores from the two forms to increase the sample size. Table 5 shows the mean and standard deviation for the mathematics test for the different test booklets. As the data in Table 5 show, the mean mathematics score varies considerably across the test versions for both grades 5 and 8. For example, grade 5 students taking the bilingual test booklets have the lowest mean in mathematics (M=43.70, SD=16.92) compared to students taking the linguistically modified test in large print (M=54.04, SD=24.96).

<table>
<thead>
<tr>
<th>States</th>
<th>Grade 5</th>
<th></th>
<th>Grade 8</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>1. Original</td>
<td>49.48</td>
<td>20.56</td>
<td>250</td>
<td>48.65</td>
</tr>
<tr>
<td>2. Linguistically Modified</td>
<td>51.50</td>
<td>20.50</td>
<td>151</td>
<td>50.57</td>
</tr>
<tr>
<td>3. Large Print</td>
<td>50.11</td>
<td>21.77</td>
<td>86</td>
<td>50.47</td>
</tr>
<tr>
<td>4. Large Print Modified</td>
<td>54.04</td>
<td>24.96</td>
<td>78</td>
<td>58.77</td>
</tr>
<tr>
<td>5. Bilingual</td>
<td>43.70</td>
<td>16.92</td>
<td>67</td>
<td>40.58</td>
</tr>
<tr>
<td>Total</td>
<td>50.00</td>
<td>21.06</td>
<td>632</td>
<td>50.00</td>
</tr>
</tbody>
</table>

It must be noted at this point that students taking the modified test version in large print are more heterogeneous (SD=24.96) than those taking the bilingual version (SD=16.92). The trend of data for students in grade 8 is very consistent with those presented for students in grade 5. Once again, students taking the bilingual booklet show the lowest performance (M=40.58, SD=13.69) and those taking the linguistically modified version with large print have the highest
performance (M=58.77, SD=19.98). These findings are consistent with earlier research on the bilingual version of assessments. When students are not instructed in their native language then the native language version of the questions may not help and may even act as a distractor. For example, Abedi, Hofstetter and Lord (2004) demonstrated that when language of assessment and language of instruction are not aligned then native language assessment may not provide valid assessment outcomes.

Performance by Subgroups

**ELLs**

We can now present performance differences between ELL, SWD, ELLWD and non-LEP/non-SD students. Table 6 presents the means and standard deviations for mathematics scores by students’ ELL status. As the data in Table 6 show, there is a substantial performance gap between ELL and non-ELL students. For example, in the original mathematics booklet, the mean for ELL students is 41.31 (SD=19.00) as compared to a mean of 58.11 (SD=17.99) for non-ELL students. Similarly, for the linguistically modified version the mean for ELL is 43.29 (SD=20.38) as compared to a mean of 55.35 (SD=19.31) for non-ELL students. A similar trend can be seen with other versions of the mathematics test. However, with some of the versions the performance gap between ELL and non-ELL students is lower than with other versions. For example, the performance-gap between ELL and non-ELL students (12.06) in the linguistically modified version is smaller than the performance gap between ELL and non-ELL in the original version (16.78), indicating that reducing unnecessary linguistic complexity of the assessment makes the test more accessible to ELL students.
### Table 6
Math test scores by students’ ELL status.

<table>
<thead>
<tr>
<th>States</th>
<th>Grade 5</th>
<th>Grade 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Original</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL</td>
<td>41.31</td>
<td>19.00</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>58.11</td>
<td>17.99</td>
</tr>
<tr>
<td>Total</td>
<td>48.65</td>
<td>20.32</td>
</tr>
<tr>
<td><strong>Linguistically Modified</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL</td>
<td>43.29</td>
<td>20.38</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>55.35</td>
<td>19.31</td>
</tr>
<tr>
<td>Total</td>
<td>49.82</td>
<td>20.64</td>
</tr>
<tr>
<td><strong>Large Print</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL</td>
<td>46.14</td>
<td>20.54</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>50.41</td>
<td>21.28</td>
</tr>
<tr>
<td>Total</td>
<td>48.01</td>
<td>20.83</td>
</tr>
<tr>
<td><strong>Linguistically Modified with</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Large Print</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL</td>
<td>37.62</td>
<td>21.86</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>59.25</td>
<td>21.73</td>
</tr>
<tr>
<td>Total</td>
<td>49.35</td>
<td>24.18</td>
</tr>
<tr>
<td><strong>Bilingual</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL</td>
<td>43.34</td>
<td>16.79</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>67.86</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SWD**

There is also a large performance gap between SWD and non-SWD students. Students with disabilities generally perform lower than non-disabled students. However, the pattern of performance gap between SWD and non-SWD is more complex than the case for ELL students. Unlike ELL students who share a common need for linguistic assistance, students with disabilities are in a variety of subgroups with very different characteristics and quite distinct needs. Some of these students perform much higher than others within the disability subgroups. Researchers are quite limited in their ability to carefully examine assessment issues for these students due to lack of access to sufficiently large numbers of students, particularly, in the low incidence categories of disabilities. Data presented in Table 7 illustrate some of these issues.
Table 7 presents summary statistics (mean, SD and N) by students’ disability status for students in grades 5 and 8 taking the mathematics test. In addition to the summary statistics mentioned above, students’ types of disability under different forms of accommodation are reported. For example, in grade 5 the mean math score for the original booklet was 36.90 (SD=20.47) compared to a mean of 49.48 (SD=20.57) for non-SWD students with a performance difference of 12.58 score points. However, the mean for SWD in the linguistically modified version dropped substantially from 36.90 to 25.83 (SD=11.51) as compared with a mean of 51.50 (SD=20.50) for non-SWD with a performance difference of 25.67 score points. This trend is quite different from what we reported for ELL students (Table 6). The data appears to indicate that ELL students benefited from the linguistic modification of assessment, whereas the opposite trend was observed for students with disabilities which may not be an accurate description of what happened.

The main reason for such performance difference across test versions may not be due to test format or language of assessment but to initial differences between students’ assessed under
different test booklets. For example, in grade 5 students in disability categories number 1 (special learning disabilities), 2 (speech or language impairments), 4 (emotional disturbance), 9 (other health impairments) and 12 (multiple disabilities) were tested using the original test booklet, while students in categories 1 (specific learning disabilities) and 3 (mental retardation) were tested with the linguistically modified version of the test. Appropriate comparisons should include students with similar characteristics (with no major initial differences). However, this may not be possible in the case of students with disabilities due to small sample sizes for many categories of disabilities.

Table 8 presents frequency and percent of SWD students tested under different accommodations in ELDA. As the data in this table suggest, the most commonly recommended accommodation by school is *extended time*. Of the 538 SWD accommodated, 254 (47.2%) received extended time accommodation. However, since ELDA is considered an un-timed test, extended time may not be conceived as a form of accommodation for this assessment. Next to the extended time accommodation, *small group testing* was the most frequent accommodation used for SWD (used by 28.8% of students with disabilities). Other accommodations such as *preferential seating* and *clarity of directions* were also used for a large group of students.

<table>
<thead>
<tr>
<th>Accommodations</th>
<th>Accommodated</th>
<th>Non-Accommodated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Pct of Tot</td>
</tr>
<tr>
<td>Extended time</td>
<td>254</td>
<td>47.2</td>
</tr>
<tr>
<td>Frequent breaks</td>
<td>24</td>
<td>4.5</td>
</tr>
<tr>
<td>Small groups</td>
<td>155</td>
<td>28.8</td>
</tr>
<tr>
<td>Repeat Instructions</td>
<td>13</td>
<td>2.4</td>
</tr>
<tr>
<td>Clarify Directions</td>
<td>39</td>
<td>7.2</td>
</tr>
<tr>
<td>Preferential Seating</td>
<td>49</td>
<td>9.1</td>
</tr>
<tr>
<td>Note Taking</td>
<td>4</td>
<td>0.7</td>
</tr>
<tr>
<td>Read Instruction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilingual Explanation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paraphrasing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reader</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternate Testing Site?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Printed Direction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any Accommodation</td>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

Table 8
Accommodations provided to students with disabilities in ELDA (Total tested: 1,443)
To see the impact of these accommodations on the performance of SWDs, we compared the performance of accommodated SWD students with the performance of non-SWD students. Table 9 presents a summary of analyses for this comparison. While this may not provide a clear picture of how effective the accommodations are in reducing the performance-gap between SWD and non-SWD students, it may be an indication of the magnitude of this gap. For example, the performance-gap between SWD and non-SWD ranged from 0.88 under the *preferential seating* accommodation to 5.55 under the *frequent breaks* accommodation. One possible explanation for this difference is that the preferential seating accommodation may be more effective in reducing the performance gap between SWD and non-SWD student. Another possible explanation might be that students under the preferential seating accommodation were higher performing initially with less performance gap between SWDs and non-SWDs than those under the frequent breaks accommodation. Later in this section, we will provide evidence on the effectiveness of accommodations in reducing the performance gap between SWD and non-SWD and on the validity of accommodated assessments for these students.

Table 9
Reading scores by student disabilities status in ELDA (Reading Subscale)

<table>
<thead>
<tr>
<th>Accommodation</th>
<th>SD M</th>
<th>SD</th>
<th>N</th>
<th>Non-SD M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended time</td>
<td>33.22</td>
<td>14.73</td>
<td>235</td>
<td>36.97</td>
<td>14.38</td>
<td>1112</td>
</tr>
<tr>
<td>Frequent breaks</td>
<td>30.85</td>
<td>20.90</td>
<td>20</td>
<td>36.40</td>
<td>14.39</td>
<td>1327</td>
</tr>
<tr>
<td>Small groups</td>
<td>32.05</td>
<td>15.69</td>
<td>146</td>
<td>36.84</td>
<td>14.28</td>
<td>1201</td>
</tr>
<tr>
<td>Repeat Instructions</td>
<td>31.33</td>
<td>18.96</td>
<td>12</td>
<td>36.36</td>
<td>14.46</td>
<td>1335</td>
</tr>
<tr>
<td>Clarify Directions</td>
<td>32.20</td>
<td>14.55</td>
<td>35</td>
<td>36.43</td>
<td>14.50</td>
<td>1312</td>
</tr>
<tr>
<td>Preferential Seating</td>
<td>35.47</td>
<td>12.88</td>
<td>47</td>
<td>36.35</td>
<td>14.57</td>
<td>1300</td>
</tr>
<tr>
<td>Note Taking</td>
<td>32.75</td>
<td>12.76</td>
<td>4</td>
<td>36.33</td>
<td>14.52</td>
<td>1343</td>
</tr>
<tr>
<td>Read Instruction</td>
<td>31.33</td>
<td>18.99</td>
<td>12</td>
<td>36.36</td>
<td>14.46</td>
<td>1335</td>
</tr>
<tr>
<td>Bilingual Explanation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paraphrasing Reader</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternate Testing Site?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Printed Direction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any Accommodation</td>
<td>32.85</td>
<td>14.21</td>
<td>346</td>
<td>37.52</td>
<td>14.42</td>
<td>1001</td>
</tr>
</tbody>
</table>

55
ELLWD

As explained earlier in this report, the main purpose of this study was to examine the effectiveness and validity of accommodations for ELL students with disabilities (ELLWD). However, due to the nature of this subgroup and to logistical issues encountered in this study, we were not able to collect data from a sufficiently large number of students that would allow meaningful analyses. Table 10 presents the number of ELLWD by different categories of disabilities. As the data in Table 10 suggest, of the total 2,675 students who participated in this study, only 120 (4%) were ELL students with disabilities. Moreover, students in this small sample were not equally distributed across the 14 categories of disabilities. For example, as Table 10 shows, of the 120 students labeled as ELLWD, 109 (91%) were in the Special Learning Disability category, leaving only 11 students for all other categories of disabilities. Therefore, examining performance differences between SWD and non-SWD students under such conditions is methodologically challenging. Unfortunately, this places a considerable limitation on our ability to reach a conclusion regarding the effectiveness and validity of accommodated assessments for these students.

Table 10
Number of ELL students with different types of disabilities in math and ELDA

<table>
<thead>
<tr>
<th>Type of Disability</th>
<th>Math</th>
<th>ELDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Specific Learning Disabilities</td>
<td>68</td>
<td>41</td>
</tr>
<tr>
<td>2) Speech/Language Impairments</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>3) Mental Retardation</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4) Emotional Disturbance</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5) Deaf/Blindness</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6) Visual Impairment</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7) Hearing Impairments</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8) Physical Impairments</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9) Other Health Impairments</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10) Autism</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11) Traumatic Brain Injury</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12) Multiple Disabilities</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13) Developmental Delays</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14) Other</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>77</td>
<td>43</td>
</tr>
</tbody>
</table>

This is particularly challenging since the main purpose of this study was on the effectiveness and validity of accommodations for these students. However, findings from students with disability and from ELL samples would shed light on the accommodated assessment for ELLWD. Accommodations that are used to address disability issues but are not
effective and valid for SWD students will similarly not be useful for ELLWD students. Similarly, accommodations that are supposed to provide linguistic assistance to ELL students but are shown to not be effective or valid for ELL students will not be useful for these students as well.

**Effectiveness of Accommodations Used in This Study**

As discussed in the design section above, an accommodation can be considered effective if it increases the performance of the recipient of the accommodation. For example, an accommodation for a student with a disability is effective if it increases the performance of that student when compared with another SWD student under the standard testing condition with no accommodation provided. Testing effectiveness of accommodations is complex and demands special arrangement since it requires students who are in need of accommodations but are not receiving those accommodations. To test the effectiveness of accommodations used in this study, we sent an assent form to parents of children who might need accommodations to know that their children might or might not receive the accommodation (Appendix F). To plan for a control group in advance, we asked teachers of the selected classrooms to provide the necessary information about all students particularly those in need of accommodations through a Class Roster. Based on the information provided by the Class Roster, we identified pairs of students with similar disability status and similar backgrounds. We randomly assigned one of these students to the accommodated testing condition and one to the standard testing condition with no accommodations provided. While this approach reduced the number of students receiving accommodation by half, it was necessary to develop a sound experimental research methodology for testing effectiveness of accommodations.

Table 11 present the results of analyses for effectiveness of all accommodations used for ELDA focusing on the results for four of the most commonly used accommodations. These accommodations include 1) extended time, 2) preferential seating, 3) small group testing, and 4) clarifying directions. To test the effectiveness of accommodations used in ELDA, we compared the performance of accommodated SWDs with the non-accommodated SWDs who had similar type of disabilities and similar background characteristics and performance indicators. We used the analysis of variance (ANOVA) controlling for the Type I error rate due to multiple
comparisons. In addition to reporting the F-ratio in Table 11, we also reported the index for the strength of association ($\omega^2$). As the data in Table 10 shows, none of the F-ratios were significant, suggesting that the accommodations did not have large enough impact on the outcome of the accommodated assessment to be noticeable.

Table 11
Effectiveness of accommodations for reading subscale of ELDA with no covariate

<table>
<thead>
<tr>
<th>Accommodation</th>
<th>Disability Status</th>
<th>Mean</th>
<th>SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>All accommodations combined</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>32.60</td>
<td>12.26</td>
<td>141</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>31.72</td>
<td>13.76</td>
<td>145</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .0.185</td>
<td>P = .668</td>
<td>$\omega^2 = .001$</td>
</tr>
<tr>
<td>Extended Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>32.37</td>
<td>13.94</td>
<td>117</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>31.48</td>
<td>12.43</td>
<td>86</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .222</td>
<td>P = .638</td>
<td>$\omega^2 = .001$</td>
</tr>
<tr>
<td>Preferential Seating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>35.17</td>
<td>11.83</td>
<td>23</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>31.58</td>
<td>13.45</td>
<td>180</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .1.49</td>
<td>P = .224</td>
<td>$\omega^2 = .007$</td>
</tr>
<tr>
<td>Small Group Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>32.77</td>
<td>14.77</td>
<td>73</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>31.55</td>
<td>12.43</td>
<td>130</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .388</td>
<td>P = .534</td>
<td>$\omega^2 = .002$</td>
</tr>
<tr>
<td>Clarifying Directions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>25.40</td>
<td>10.11</td>
<td>15</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>32.52</td>
<td>13.40</td>
<td>188</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .4.04</td>
<td>P = .046</td>
<td>$\omega^2 = .020$</td>
</tr>
</tbody>
</table>

However, an interesting trend emerged in the effectiveness of accommodations in this study. In four of the five cases, the accommodated group performed better than the non-accommodated group, although the sample size was not large enough to detect that difference. When we used students’ state reading test scores as a covariate (Table 12), the trend of effectiveness changed with some accommodations, but once again, there is no statistical power in the comparison due to the small sample size.
Table 12
Effectiveness of accommodations for reading subscale of ELDA with state reading test score as a covariate

<table>
<thead>
<tr>
<th>Accommodation</th>
<th>Disability Status</th>
<th>Mean</th>
<th>SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>All accommodations combined</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>32.48</td>
<td>14.11</td>
<td>97</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>33.56</td>
<td>12.51</td>
<td>109</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .170</td>
<td>P=.681</td>
<td>$\omega^2$ = .001</td>
</tr>
<tr>
<td>Extended Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>33.30</td>
<td>14.41</td>
<td>88</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>32.09</td>
<td>12.58</td>
<td>76</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .216</td>
<td>P=.643</td>
<td>$\omega^2$ = .001</td>
</tr>
<tr>
<td>Preferential Seating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>35.17</td>
<td>11.83</td>
<td>23</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>32.34</td>
<td>13.82</td>
<td>141</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .928</td>
<td>P=.337</td>
<td>$\omega^2$ = .006</td>
</tr>
<tr>
<td>Small Group Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>34.44</td>
<td>15.27</td>
<td>52</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>31.95</td>
<td>12.69</td>
<td>112</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = 2.81</td>
<td>P=.096</td>
<td>$\omega^2$ = .017</td>
</tr>
<tr>
<td>Clarifying Directions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>SWD</td>
<td>25.10</td>
<td>9.74</td>
<td>10</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>SWD</td>
<td>33.23</td>
<td>13.65</td>
<td>154</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = 2.57</td>
<td>P=.111</td>
<td>$\omega^2$ = .016</td>
</tr>
</tbody>
</table>

Validity of Assessment under the Accommodations Used in This Study

As with the effectiveness of accommodations, examining the validity of accommodations required a careful design and complex testing conditions. In this study, we examined the validity of accommodations using two different approaches: 1) performance difference between accommodated and non-accommodated testing conditions, and 2) testing the invariance of structural relationships between accommodated and non-accommodated assessments.

1. Performance differences

Accommodations are supposed to help those who are in need of them without having any impact on the outcome of assessment for others who are not the intended target of the
accommodations. That is, an accommodation provides a valid outcome if it does not alter the construct being measured. To examine the unintended impact of accommodations, we tested non-ELL/non-SWD students under both accommodated and non-accommodated conditions. After receiving student background data from teachers through the Class Roster, we identified the groups to be tested for the validity purposes. For each accommodated student, we selected a match from the non-ELL/non-SWD population (called “twin” in this study) with similar background and academic performance. The twins, who under the standard testing condition would not receive accommodation, were also accommodated. Comparing the performance outcome of twins with a control group similar to the twins in many important aspects would shed light on the validity of accommodated assessments.

Table 13
Validity of accommodations for reading subscale of ELDA

<table>
<thead>
<tr>
<th>Accommodation</th>
<th>Disability Status</th>
<th>Mean</th>
<th>SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>All accommodations combined</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>Non-SWD</td>
<td>36.63</td>
<td>12.78</td>
<td>123</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>Non-SWD</td>
<td>35.21</td>
<td>13.49</td>
<td>123</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .712</td>
<td>P = .399</td>
<td>$\omega^2 = .003$</td>
</tr>
<tr>
<td>Extended Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>Non-SWD</td>
<td>35.89</td>
<td>13.07</td>
<td>98</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>Non-SWD</td>
<td>35.94</td>
<td>13.22</td>
<td>148</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .011</td>
<td>P = .976</td>
<td>$\omega^2 = .000$</td>
</tr>
<tr>
<td>Frequent Breaks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>Non-SWD</td>
<td>42.13</td>
<td>16.40</td>
<td>8</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>Non-SWD</td>
<td>35.71</td>
<td>13.00</td>
<td>238</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .1.85</td>
<td>P = .175</td>
<td>$\omega^2 = .008$</td>
</tr>
<tr>
<td>Preferential Seating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>Non-SWD</td>
<td>36.38</td>
<td>13.22</td>
<td>21</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>Non-SWD</td>
<td>35.88</td>
<td>13.16</td>
<td>225</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .0.028</td>
<td>P = .866</td>
<td>$\omega^2 = .000$</td>
</tr>
<tr>
<td>Small Group Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>Non-SWD</td>
<td>34.53</td>
<td>13.86</td>
<td>58</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>Non-SWD</td>
<td>36.35</td>
<td>12.91</td>
<td>188</td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td>F = .842</td>
<td>P = .360</td>
<td>$\omega^2 = .003$</td>
</tr>
<tr>
<td>Clarifying Directions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accommodated</td>
<td>Non-SWD</td>
<td>35.87</td>
<td>12.06</td>
<td>15</td>
</tr>
<tr>
<td>Non-Accommodated</td>
<td>Non-SWD</td>
<td>35.92</td>
<td>13.23</td>
<td>231</td>
</tr>
</tbody>
</table>

Table 13 presents the findings of the validity studies for ELDA. The accommodated and
non-accommodated assessment outcomes were compared using an analysis of variance (ANOVA) model. In the analyses of variance, adjustments were made to control for the Type I error rates due to the multiple comparisons. As Table 13 shows, none of the differences between the accommodated and non-accommodated assessment were statistically significant. These results suggest that the accommodations used in this study did not alter the construct being measured.

2. **Structural Invariance**

A multiple-group confirmatory factor analysis model was applied to the data in both mathematics and ELDA assessments. Figure 2 shows the results for grade 5 mathematics. For the mathematics assessment the 30 mathematics items were divided into four parcels with two parcels containing 8 items and two with 7 items. The homogeneity within and between parcels was carefully reviewed (see, Abedi, 1997 for issues concerning parceling test items).

![Multiple-group confirmatory factor analysis model](image)

Invariance of factor loadings of the measured variable (the four mathematic parcels), and the mathematic latent variable (Factor 1) was tested between accommodated and non-accommodated assessments. Table 14 presents results of multiple-group factor analyses for mathematics. As data in Table 14 shows, the factor Loadings of the four mathematic parcels with
the mathematic latent variable (F1) is almost identical across the accommodation conditions (accommodated versus non-accommodated). For example, factor loadings for parcel 1 through 4 under accommodated condition are: .801, .846, .820, and .819 as compared with the respected factor loading of .816, .826, .866, .790 under non-accommodated conditions. These factor loadings are almost identical across the accommodation conditions. Consequently, the communalities (proportion of the variance explained by the factor) remained the same for the mathematic assessment under accommodated and non-accommodated conditions. We tested the invariance between factor loadings under the accommodated and non-accommodated conditions. The results confirm the invariance between the factor loadings under the two conditions.

Table 14
Multiple group confirmatory factor analyses results for Math by accommodation status

<table>
<thead>
<tr>
<th>Description / Test</th>
<th>Accommodated</th>
<th>Non-Accommodated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Par 1</td>
<td>Par 2</td>
</tr>
<tr>
<td>Factor</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Factor Loadings</td>
<td>.801</td>
<td>.846</td>
</tr>
<tr>
<td>Error</td>
<td>.598</td>
<td>.534</td>
</tr>
<tr>
<td>Communality</td>
<td>.642</td>
<td>.715</td>
</tr>
</tbody>
</table>

Similarly, for ELDA reading and writing scores were used as the measured variables to create an English language proficiency (ELP) latent variable using a multiple group confirmatory factor analytic model (Figure 3). Table 15 presents a summary of results for this model.

Figure 3. Multiple-group confirmatory factor analysis model for testing invariance between accommodated and non-accommodated assessments in ELDA.
As the data in Table 15 shows, the factor loadings of reading and writing with the English proficiency latent variable (F1) are very similar across the accommodated and non-accommodated conditions. For example, the correlation between ELDA reading and ELP latent variable is .729 under accommodation condition as compared with a correlation of .722 under non-accommodation condition. For writing, the correlation with the ELP latent variable under accommodation is .92 as compared with the correlation of .90 for the non-accommodated assessment. Invariance between factor structure of accommodated and non-accommodated assessment did not show an significant results therefore, the hypothesis of invariance between accommodated and non-accommodated ELDA was confirmed.

<table>
<thead>
<tr>
<th>Description / Test</th>
<th>Accommodated</th>
<th>Non-Accommodated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description / Test</td>
<td>Accommodated</td>
<td>Non-Accommodated</td>
</tr>
<tr>
<td>Reading</td>
<td>.729</td>
<td>.722</td>
</tr>
<tr>
<td>Writing</td>
<td>.920</td>
<td>.900</td>
</tr>
<tr>
<td>Error</td>
<td>.685</td>
<td>.392</td>
</tr>
<tr>
<td>Error</td>
<td>.392</td>
<td>.692</td>
</tr>
<tr>
<td>Error</td>
<td>.692</td>
<td>.436</td>
</tr>
<tr>
<td>Communality</td>
<td>.531</td>
<td>.521</td>
</tr>
<tr>
<td>Communality</td>
<td>.846</td>
<td>.810</td>
</tr>
</tbody>
</table>

As the data in Tables 13 and 14 suggest, the accommodated and non-accommodated assessments produced the same structural outcomes.

3. Validity of Accommodation Using External Criteria

As discussed earlier, in addition to the data from the math and ELDA assessments, we also collected state test scores on math and reading and student and teacher background data
through the background questionnaires. The purpose of obtaining these data was to use them as external criteria for examining the validity of accommodated assessments. In the background questionnaire, students were asked to provide some background information such as gender, ethnicity, country of origin, number of years in the U.S. and amount of a language other than English at home. In addition to the background questions, students were asked to self report their proficiencies in their native language as well as in English in four modalities (reading, writing, speaking and understanding) using a 5-point Likert-Scale (1 little or no proficiency to 5, high level of proficiency).

In this study we used both sets of variables, the state test scores in math and reading and also self-reported academic information (L1 and L2 proficiencies, how often studied subjects in other languages, number of times changed schools, and how far in school). We created two multiple regression models, Model 1 and Model 2. In both models the total test score (ELDA total NCE mathemetic score) was used as the criterion variable. In Model 1 state reading and mathematic test scores were used as the predictors and in Model 2 the background variables were used as predictors. These background variables were all in Likert-scale. We applied the two models on the data from students with disabilities (SWD) who were accommodated and to non-SWDS/non-accommodated students as a control group. The results of analyses for the two content areas (ELDA and mathematics) were very similar; therefore, we report the results for mathematics only.

Table 16
Multiple regression results, SWD, accommodated students in grades 5 and 8 (Math NCE as the criterion and state test scores as predictors)

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficient b</th>
<th>SEb</th>
<th>β</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>41.53</td>
<td>2.70</td>
<td></td>
<td>15.39</td>
<td>0.000</td>
</tr>
<tr>
<td>State Math</td>
<td>6.75</td>
<td>4.35</td>
<td>0.27</td>
<td>1.55</td>
<td>0.127</td>
</tr>
<tr>
<td>State Reading</td>
<td>4.37</td>
<td>4.16</td>
<td>0.19</td>
<td>1.05</td>
<td>.298</td>
</tr>
<tr>
<td>Summary Statistics</td>
<td>R²=0.180</td>
<td>F=5.60</td>
<td>p=0.006</td>
<td>n=54</td>
<td></td>
</tr>
</tbody>
</table>

Table 16 presents a summary of multiple regression analyses for Model 1 applied to the data on accommodated SWD students and Table 17 presents similar results for Model 1 applied to non-
SWD/non-accommodated students. Although the number of subjects differ substantially across the SWD (n=54) and non-SWD (n=752) but there are some similarities in the trend of outcomes for the two groups of students. For example, for the SWD students (Table 16), state math had higher predictive power ($\beta = 0.27$) than state reading ($\beta = 0.19$). Similar pattern was found for the non-SWD students (Table 17). For these students state math had higher impact on the prediction ($\beta = 0.67$) than state reading ($\beta = 0.13$).

Table 17
Multiple regression results, Non-SWD, non-accommodated students in grades 5 and 8 (Math NCE as the criterion and state test scores as predictors)

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficient</th>
<th>Standardized Coefficient</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>SE_b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>50.54</td>
<td>.610</td>
<td>82.89</td>
<td>0.000</td>
</tr>
<tr>
<td>State Math</td>
<td>12.82</td>
<td>1.27</td>
<td>10.14</td>
<td>0.000</td>
</tr>
<tr>
<td>State Reading</td>
<td>2.51</td>
<td>1.27</td>
<td>1.97</td>
<td>0.049</td>
</tr>
<tr>
<td>Summary Statistics</td>
<td>$R^2=0.308$</td>
<td>F=166.49</td>
<td>p=0.000</td>
<td>n=752</td>
</tr>
</tbody>
</table>

The model produced significant results for both SWD (F=5.60, p=.000) and non-SWD (F = 166.49, p=0.000). However, the $R^2$ for SWDs (0.180) is substantially lower than the $R^2$ for non-SWD (0.308). This difference could be explained partly by differences in the sample size and sample characteristics across the SWD/non-SWD samples.

Table 18 summarizes the results of analyses for Model 2 (using background variables as predictors) when applied to accommodated SWDs and Table 19 presents similar results for non-accommodated non-SWDs. Based on the data from Table 18, students’ background variables and their self-reported language proficiency were powerful predictors of student performance in math ($R^2 = .425$). For non-SWDs, however, the power of background variables in predicting math score ($R^2 = .103$) was not as strong as for SWDs. Once again, this may partly explained by differences in the sample characteristics.
Table 18
Multiple regression results, SWD, accommodated students in grades 5 and 8 (Math NCE as the criterion and student background variables as predictors)

<table>
<thead>
<tr>
<th>Description</th>
<th>Unstandardized Coefficient</th>
<th>Standardized Coefficient</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>66.33</td>
<td>18.87</td>
<td>3.52</td>
<td>0.001</td>
</tr>
<tr>
<td>How often a non-English speak a non-English lang</td>
<td>-0.17</td>
<td>2.97</td>
<td>-0.01</td>
<td>-0.06</td>
</tr>
<tr>
<td>Read a non-English lang</td>
<td>0.08</td>
<td>3.43</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Understand a non-English lang</td>
<td>5.32</td>
<td>4.10</td>
<td>0.28</td>
<td>1.30</td>
</tr>
<tr>
<td>Write a non-English lang</td>
<td>2.81</td>
<td>4.10</td>
<td>0.16</td>
<td>0.69</td>
</tr>
<tr>
<td>Studied math in another language</td>
<td>-6.79</td>
<td>3.84</td>
<td>-0.37</td>
<td>-1.77</td>
</tr>
<tr>
<td>Studied any subjects in another language</td>
<td>-7.70</td>
<td>3.84</td>
<td>-0.42</td>
<td>-2.00</td>
</tr>
<tr>
<td>How well speak English</td>
<td>4.53</td>
<td>6.64</td>
<td>0.18</td>
<td>0.68</td>
</tr>
<tr>
<td>How well read English</td>
<td>-5.06</td>
<td>9.34</td>
<td>-0.21</td>
<td>-0.54</td>
</tr>
<tr>
<td>How understand English</td>
<td>-7.33</td>
<td>7.40</td>
<td>-0.27</td>
<td>-0.99</td>
</tr>
<tr>
<td>How well write English</td>
<td>1.05</td>
<td>9.15</td>
<td>0.04</td>
<td>0.12</td>
</tr>
<tr>
<td>Times changed school</td>
<td>2.06</td>
<td>1.92</td>
<td>0.16</td>
<td>1.07</td>
</tr>
<tr>
<td>How far in school</td>
<td>-2.22</td>
<td>1.50</td>
<td>-0.22</td>
<td>-1.48</td>
</tr>
</tbody>
</table>

Summary Statistics: R^2=0.425  F=1.99  p=0.053  n=49

As it is evident from the overall R^2 the background variables and self-reported proficiencies showed strong prediction power for SWDs than for the non-SWD subgroup. Table 19 presents a summary of regression analysis report for Model 2 for non-SWD students. As the data in Table 19 show, none of the 13 predictors had significant power in predicting students’ performance in math. Comparing these findings for non-SWDs with those reported for SWDs reveals inconsistencies between the regression outcomes for the two groups. As indicated earlier, one of the most likely explanations for such inconsistencies might be differences between the sample characteristics for the two groups. For example, it is clear that students in the non-SWD group have shown higher academic performance in general. Therefore, the two groups of students may not be initially different.
Table 19
Multiple regression results, Non-SWD, non-accommodated students in grades 5 and 8 (Math NCE as the criterion and student background variables as predictors)

<table>
<thead>
<tr>
<th>Model Description</th>
<th>Unstandardized Coefficient</th>
<th>Standardized Coefficient</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>56.82</td>
<td>4.03</td>
<td>14.10</td>
<td>0.000</td>
</tr>
<tr>
<td>How often a non-English</td>
<td>1.77</td>
<td>.83</td>
<td>2.13</td>
<td>0.034</td>
</tr>
<tr>
<td>Speak a non-English lang</td>
<td>1.66</td>
<td>1.36</td>
<td>1.22</td>
<td>0.224</td>
</tr>
<tr>
<td>Read a non-English lang</td>
<td>-1.22</td>
<td>1.30</td>
<td>-0.94</td>
<td>0.349</td>
</tr>
<tr>
<td>Understand a non-English lang</td>
<td>-0.19</td>
<td>1.21</td>
<td>-0.16</td>
<td>0.873</td>
</tr>
<tr>
<td>Write a non-English lang</td>
<td>0.21</td>
<td>1.24</td>
<td>0.17</td>
<td>0.869</td>
</tr>
<tr>
<td>Studied math in another language</td>
<td>-0.46</td>
<td>0.83</td>
<td>-0.55</td>
<td>0.586</td>
</tr>
<tr>
<td>Studied any subjects in another language</td>
<td>-0.46</td>
<td>0.92</td>
<td>-0.50</td>
<td>0.616</td>
</tr>
<tr>
<td>How well speak English</td>
<td>0.89</td>
<td>1.84</td>
<td>0.49</td>
<td>0.628</td>
</tr>
<tr>
<td>How well read English</td>
<td>-6.91</td>
<td>1.83</td>
<td>-3.77</td>
<td>0.000</td>
</tr>
<tr>
<td>How understand English</td>
<td>-0.09</td>
<td>2.14</td>
<td>-0.04</td>
<td>0.968</td>
</tr>
<tr>
<td>How well write English</td>
<td>-0.91</td>
<td>1.86</td>
<td>-0.49</td>
<td>0.624</td>
</tr>
<tr>
<td>Times changed school</td>
<td>-0.80</td>
<td>0.63</td>
<td>-1.27</td>
<td>0.203</td>
</tr>
<tr>
<td>How far in school</td>
<td>0.64</td>
<td>0.49</td>
<td>1.30</td>
<td>0.195</td>
</tr>
</tbody>
</table>

Summary Statistics

$R^2 = 0.103$  $F = 6.03$  $p = 0.000$  $n = 693$

To test the hypothesis of impact of students’ family and academic background on the outcomes of these analyses, we used the propensity score matching methodology (see for example, Foster, 2003) to select a control group from non-SWD students with similar background. We used student scores in state reading and math as the matching variables. We selected a group of about the same size ($n = 44$) as of the SWD group with similar level of academic performance.

Table 20 presents the results of regression analyses for Model 2 for the new control group based on the propensity approach. The value of $R^2$ increased from .103 with the full sample of non-SWD students to .266 which is closer to the $R^2$ of .425 for the SWD students. However, the pattern of prediction power of individual variables as evident in the standardized coefficient column is quite different across the SWD/non-SWD students.
Table 20
Multiple regression results, SWD, accommodated students in grades 5 and 8 using Propensity matched model (Math NCE as the criterion and student background variables as predictors)

<table>
<thead>
<tr>
<th>Description</th>
<th>Unstandardized Coefficient</th>
<th>Standardized Coefficient</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>44.96</td>
<td>10.44</td>
<td>4.31</td>
<td>0.000</td>
</tr>
<tr>
<td>How often a non-English Speak</td>
<td>-4.78</td>
<td>3.50</td>
<td>-1.37</td>
<td>0.182</td>
</tr>
<tr>
<td>Speak a non-English lang</td>
<td>8.96</td>
<td>5.97</td>
<td>1.50</td>
<td>0.144</td>
</tr>
<tr>
<td>Read a non-English lang</td>
<td>-3.68</td>
<td>4.22</td>
<td>-0.87</td>
<td>0.390</td>
</tr>
<tr>
<td>Understand a non-English lang</td>
<td>-3.70</td>
<td>4.53</td>
<td>-0.82</td>
<td>0.420</td>
</tr>
<tr>
<td>Write a non-English lang</td>
<td>0.58</td>
<td>4.72</td>
<td>0.12</td>
<td>0.903</td>
</tr>
<tr>
<td>Studied math in another language</td>
<td>1.85</td>
<td>2.35</td>
<td>0.79</td>
<td>0.436</td>
</tr>
<tr>
<td>Studied any subjects in another lang</td>
<td>-5.18</td>
<td>2.64</td>
<td>-1.96</td>
<td>0.059</td>
</tr>
<tr>
<td>How well speak English</td>
<td>-0.40</td>
<td>7.65</td>
<td>-0.05</td>
<td>0.959</td>
</tr>
<tr>
<td>How well read English</td>
<td>2.23</td>
<td>6.59</td>
<td>0.34</td>
<td>0.738</td>
</tr>
<tr>
<td>How understand English</td>
<td>-0.91</td>
<td>6.85</td>
<td>-0.13</td>
<td>0.896</td>
</tr>
<tr>
<td>How well write English</td>
<td>-5.95</td>
<td>8.74</td>
<td>-0.68</td>
<td>0.501</td>
</tr>
<tr>
<td>Times changed school</td>
<td>-2.41</td>
<td>2.28</td>
<td>-1.06</td>
<td>0.298</td>
</tr>
<tr>
<td>How far in school</td>
<td>1.33</td>
<td>1.25</td>
<td>1.06</td>
<td>0.297</td>
</tr>
</tbody>
</table>

Summary Statistics
\[ R^2 = 0.266 \quad F = 0.84 \quad p = 0.622 \quad n = 44 \]

The results of these analyses suggest that when accommodated and non-accommodated assessments are compared with more reliable and more objective criteria of academic performance (state test scores in math and reading), then more consistency between the accommodated and non-accommodated assessments can be seen but when less reliable and less objective criteria, such as self-reported proficiency are used, the inconsistencies are quite apparent. This trend of finding is expected since a major requirement in criterion validation approach is that the criterion variables must be quite reliable and valid.

In general, the results of the criterion-related validity approach confirm the findings from the “group differences” approach and suggest that the accommodated and non-accommodated assessments are equally valid.
Discussion

Federal and state legislations mandate inclusion of all students in national and state assessments including special needs student population (ELLs, SWDs and ELLWDs). However, results of many studies conducted recently suggest that these students lag far behind their peers in all content areas (see, for example, Abedi, Hofstetter & Lord, 2004, Solano-Flores, & Li, 2006; Thurlow, Elliott, & Ysseldyke, 2003; Tindal, Anderson, Helwig, Miller & Glasgow, 2000). Thus, mandating inclusion of these students without attention to factors affecting their academic performance may not be productive and may even put additional burden on these students, their parents and their teachers. To provide equal education opportunity for these students, and to expect them to function well within the federal and state accountability systems, issues concerning instruction and assessment of these students should be carefully examined. Many different factors may impact the academic performance of these students including factors related to both instruction and assessment. For example, findings from recent studies have suggested that assessment outcomes for these students may be affected by different nuisance or extraneous variables that may invalidate the results of assessments for these students. To reduce the impact of such nuisance variables, researchers, measurement experts and educational practitioners suggest the use of accommodation to “level the playing field”.

However, research on the equity of assessments has uncovered major issues regarding the accommodations used for these students. Among these issues are concerns regarding the effectiveness, validity, differential impact and feasibility of accommodations used for ELLs, SWDs, and ELLWDs. For example, as discussed earlier in this report, a recent review of accommodations currently used by states reveal that many of these accommodations may not be appropriate, they may not be effective, and they may produce invalid assessment outcomes. On the other hand examining the impact of such issues on the assessment and accommodations of special needs student population (ELL, SWD and ELLWD) is not an easy task; it requires carefully designed studies and testing large numbers of students. Obtaining large enough number of subjects for examining the effectiveness and validity of accommodations is a challenging task for all three subgroups particularly for ELLWD students. Another major obstacle in the study of accommodations is the variety of accommodations that are being used by states. Carefully
examining different aspects of these accommodations also is a big challenge.

Among the many different accommodations used by different states (Rivera, et al, 2006), we used two different sets of accommodations for ELL, SWD and ELLWD students: those recommended by schools based on students’ IEP or 504 plans and some of those from the list of research supported accommodations. We examined the effectiveness and validity of these accommodations by comparing the performance of different groups of students. For examining the effectiveness of accommodations, we compared the performance of ELL/SWD/ELLWD students tested under accommodations with a similar groups of students tested under the standard testing condition with no accommodations provided. To examine the validity of accommodations, we compared the performance of non-ELL/non-SWD students under accommodated and non-accommodated conditions. Additionally, we tested the validity of accommodations through a criterion-validation approach. We used state test scores and teacher rating of student performance as external criteria. We compared accommodated and non-accommodated outcomes with the external criteria through testing invariance of structural relationship between the assessment outcomes (both accommodated and non-accommodated) using a multiple group confirmatory factor analysis model. The results of validation studies using both approaches (comparisons between accommodated and non-accommodated students and criterion-related validation) consistently suggested that many of the accommodations used in this study that are state approved are valid. These findings provide enough evidence to justify aggregation of the accommodated and non-accommodation assessments.

The results of the effectiveness studies showed a trend of increasing performance of students under the accommodated assessments, but due to the small sample size, the effectiveness of none of these accommodations used in this study reached a level of statistical significance. Conversely, the results of our validity studies clearly suggest that the accommodations used in this study did not compromise the validity of assessment. The results of multiple group factor analyses also confirmed that the provision of accommodations did not affect the validity of assessments.

The outcome of this study has two major implications for the assessment of the special needs populations. First, this study revealed substantial challenges for comprehensive studies
focusing on the assessment and accommodations for the three subgroups of special needs students (ELLs, SWDs, and ELLWDs) who lag far behind their peers. The ultimate goal of this study which included multiple sites from many locations nationwide was to have a nationally representative and large enough subjects to have power in testing effectiveness and validity of accommodations particularly for students in the low incident disability categories. However, in spite of the best efforts by the project team, it proved to be extremely difficult to secure enough testing sites to provide sufficient data for a careful and comprehensive review of accommodation issues. Therefore, future studies focusing on these subgroups of students must be cognizant of these issues. Second, given the logistical difficulty, the outcome of this study provided evidence suggesting that many of the state-supported accommodations do not alter the construct being measured. Therefore, their use in the state assessment and accountability systems can be permitted. The findings of this study reveal the methodological issues in the assessment and accountability of special needs student population and point to the need for more attention to accommodation issues for these students.

**Limitations of This Study**

Among the special needs student populations, ELL students with disabilities suffer the most in their academic career as they face two challenges, learning a second language and overcoming their disabilities. The purpose of this study was to provide information on the effectiveness and validity of accommodations that can help to obtain more valid academic assessments for these students. As researchers, we were faced with several challenges, the most important of which was finding sufficiently large numbers of students to test the various technical aspects of accommodations for these students. The problem of sample size became even more serious for students in the low incidence disability categories.

A major problem impacting the sample size in this study was the lack of cooperation from many of the states that initially had agreed to participate. Therefore, in spite of substantial efforts by the researchers and by the project team to obtain large number of students from different states, there were problems in securing assessment sites. However, by including some states that were not initially part of the project, we were able to obtain sufficiently large numbers of subjects to examine the accommodation issues, at least for ELL and SWD students.
Another problem was variation in testing time. As was elaborated in the proposal for this study, the best time for testing students was late spring because students at that time have acquired the knowledge and skills to respond to test items both in math and in ELD. However, many state denied access to their sampled classes due to priority given to school testing. Many of these schools provided access to their classrooms during the fall or early winter. This introduced another extraneous variable that impacted the outcome of this study.
References


California, National Center for Research on Evaluation, Standards, and Student Testing.


Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in


Assessment accommodations research: Considerations for design and analysis (NCEO Tech. Rep. 26). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.


Appendices
Appendix A:

Student Background Questionnaire

Student First Name: __________________________ Last Name: ________________
Teacher Name: ____________________________
School Name: _____________________________
State: ________________________________

Student will complete this questionnaire.
Student Background Questionnaire (Math)
The information you provide will be used for research purposes only and will be kept strictly confidential

Please fill in the circle (O) next to your answer:

1. In what country were you born?
   - China O
   - Korea O
   - Mexico O
   - U.S. O
   - Other O

2. In what country did you live for most of your life?
   - China O
   - Korea O
   - Mexico O
   - U.S. O
   - Other O

3. How long have you lived in the United States? ____________ Years

4. What is your birth date? _____/_____/_____
   - Month
   - day
   - year

5. What grade are you in? _______ grade

6. Are you male or female?   Male O  Female O

7. How often you speak a language other than English at home?
   - Always O
   - Most of the time O
   - Rarely O
   - Never O

8. If you speak a language other than English, how well do you …
   - speak that language?  Very well O  Well O  Not well O  Not at all O
   - read that language?    Very well O  Well O  Not well O  Not at all O
   - understand that language? Very well O  Well O  Not well O  Not at all O
   - write that language?  Very well O  Well O  Not well O  Not at all O

9. Before you started going to school, you spoke
   - English O  Chinese O  Korean O  Spanish O  Other (please specify) O

10. What is your zip code? ______________

11. Which best describes you (check one)?
   - White O
   - African American or Black O
   - Hispanic O
   - Asian O
   - Other (please specify) O

12. Have you studied mathematics in a language other than English?
13. Have you studied any subjects at school in a language other than English?

Never ○ Less than one year ○ Over a year ○ Always ○

14. How well do you?

   Speak English?          Very well ○    Well ○ Not well ○ Not at all ○
   Read English?           Very well ○    Well ○ Not well ○ Not at all ○
   Understand English?     Very well ○    Well ○ Not well ○ Not at all ○
   Write English?          Very well ○    Well ○ Not well ○ Not at all ○

15. In the last two years, how many times have you changed schools because you moved?

   Never ○ One time ○ Two times ○ Three times ○ Four or more times ○

16. How far do you think you will go in school?

   ○ I will graduate from high school
   ○ I will have some education after high school
   ○ I will graduate from college
   ○ I will go to graduate school
   ○ I will not finish high school
   ○ I don't know

Thank You
Appendix B:

Validity of Accommodations for LEP Students with Disabilities in Math and English Proficiency Content Assessment

The Oklahoma State Department of Education
Council of Chief State School Officers
Advance Research & Data Analyses Center

Teacher Questionnaire

Teacher Name: __________________________

School Name: __________________________

District: _______________________________

State: _________________________________

Teacher will complete this questionnaire.
(Please indicate the grade level on each page).
Teacher Questionnaire

The information you provide will be used for research purposes only and will be kept strictly confidential.

Please answer the following questions based on your class that is participating in this research study:

1. School Name ______________________       Teacher Name ______________________

2. Language of Instruction: ___ English only
   (Check one)       ___ Spanish only
   ___ Sheltered English (English, with native language infused)
   ___ Other (please specify)

3. How many months have you been teaching this classroom of students? ________ Months

4. How many of the students in your class are:
   a. Limited English Proficient (LEP) - non-native English speakers ______
   b. Fluent English Proficient (FEP) - originally LEP, transitioned to FEP ______
   c. Initially Fluent in English (IFE) - native English speakers or bilingual students fluent in English when first tested ______

5. How many students in your class have disabilities? ____________

6. In terms of ethnic background, how many of your students are:
   a. Latino/Hispanic ______
   b. Caucasian ______
   c. African-American ______
   d. Asian/Pacific Islander ______
   e. Other (please specify) ______

7. In terms of native language, how many of your students speak:
   a. English ______
   b. Spanish ______
   c. Vietnamese_______
   d. Chinese_______
   e. Other (please specify) ______
9. In what type of math class are your students enrolled?
   - mainstream math class
   - bilingual math class
   - gifted mainstream math class
   - gifted bilingual math class
   - this varies depending on the student
   - Other (please specify)________

10. In terms of math achievement for this grade level, how many of your students are:
   - a. below grade level ______
   - b. at grade level ______
   - c. above grade level ______

11. How many hours per week do you have designated for math teaching preparation?
   - None
   - Less than 1
   - 1 to 2
   - More than 2

12. In terms of English reading proficiency, how many of your students are:
   - a. below grade level ______
   - b. at grade level ______
   - c. above grade level ______

13. In terms of English writing proficiency, how many of your students are:
   - a. below grade level ______
   - b. at grade level ______
   - c. above grade level ______

14. In terms of English speaking proficiency, how many of your students are:
   - a. below grade level ______
   - b. at grade level ______
   - c. above grade level ______

15. Including this year, how many years in total (including part-time teaching) have you taught English language learners (ELLs)?
   - 2 years or less
   - 3-5 years
16. In what type of language arts classes are your students enrolled?

- bilingual language arts class
- mainstream language arts class
- gifted mainstream language arts class
- gifted bilingual language arts class
- this varies depending on the student
- Other (please specify)----------------------

17. Do you have teaching certification in any of the following areas that is recognized by your state? Please fill in one choice per line.

<table>
<thead>
<tr>
<th>Certification</th>
<th>Yes</th>
<th>No</th>
<th>Not offered in my state</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Subject – Math</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single Subject – ______________________________</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English as a Second Language Specialist Credential</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify): __________________________</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

18. What is your highest academic degree?

- High school diploma
- Associate’s degree/vocational certification
- Bachelor’s degree
- Master’s degree
- Education specialist/professional diploma based on at least one year’s work past master’s degree
- Doctorate
- Professional degree (e.g., M.D., LL.B., J.D., D.D.S.)

19. What were your major fields of study in college? Please fill in all choices that apply.

- Education
- English
- History
- Math
- Other
- No graduate-level study
20. What is your gender?

- Male
- Female

21. Which best describes you?

- White (not of Hispanic origin)
- Black or African American (not of Hispanic origin)
- Hispanic
- Asian
- Pacific Islander
- American Indian or Alaskan Native
- Native Hawaiian or Other Pacific Islands
- Other (please specify)

22. Including this year, how many years in total (include part-time teaching) have you taught at the elementary/secondary level?

- 2 years or less
- 3-5 years
- 6-10 years
- 11-24 years
- 25 years or more
Accommodations for ELL students
23. In Table A below, please list specific accommodations in this class provided to your English language learner students in instruction, classroom testing, and large-scale assessments.

Table A. Accommodations used for ELL students

<table>
<thead>
<tr>
<th>Accommodations</th>
<th>Used in Instruction</th>
<th>Classroom Testing</th>
<th>Large-Scale Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

Accommodations for students with disabilities
24. In Table B below, please list specific accommodations provided to your students with disabilities in instruction, classroom testing, and large-scale assessments.

Table B. Accommodations used for students with disabilities

<table>
<thead>
<tr>
<th>Accommodations</th>
<th>Used in Instruction</th>
<th>Classroom Testing</th>
<th>Large-Scale Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

Thank You
Appendix C:

Validity of Accommodations for LEP Students with Disabilities in Math and English Proficiency Content Assessment

The Oklahoma State Department of Education
Council of Chief State School Officers
Advance Research & Data Analyses Center

School Questionnaire

Principal Name: _________________________
School Name: __________________________
District: ______________________________
State: ________________________________

School Principal will complete this questionnaire.
School Questionnaire

The information you provide will be used for research purposes only and will be kept strictly confidential.

Name: ____________________________
Title: ____________________________
School Name: ______________________________________________
Date: ___________________

1. What percentage of students in your school are:
   a. Limited English Proficient (LEP) - non-native English speakers   _____%
   b. Fluent English Proficient (FEP) - originally LEP, transitioned to FEP   _____%
   c. Initially Fluent in English (IFE) - native English speakers or bilingual students fluent in English when first tested   _____%

2. What percentage of students in your school have disabilities? ____________%

3. How many books do you have in your school library? _________________

4. How many computers in your school are available for student use (lab, classrooms, etc.)? ______________

5. Do you offer tutoring? In what subject areas? _________________________________

6. Do you offer after-school programs? _________________________________

7. What percentage of students in your school are eligible for free or reduced lunch? _______

Thank You
Appendix D:

Validity of Accommodations for LEP Students with Disabilities in Math and English Proficiency Content Assessment

The Oklahoma State Department of Education
Council of Chief State School Officers
Advance Research & Data Analyses Center

Class Roster Questionnaire

Teacher Name: _________________________
School Name: _________________________
Grade Level: _________________________
District: _____________________________
State: ______________________________

Teacher will complete this questionnaire.
(Please indicate the grade level on each page).
<table>
<thead>
<tr>
<th>Student Name (Last Name, First Name)</th>
<th>Gender M/F</th>
<th>Ethnicity</th>
<th>Disability Status</th>
<th>LEP Status</th>
<th>Receiving Free-Reduced Lunch?</th>
<th>State Math Score</th>
<th>State Reading Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Name (Last Name, First Name)</td>
<td>Gender M/F</td>
<td>Ethnicity</td>
<td>Disability Status</td>
<td>LEP Status</td>
<td>Receiving Free-Reduced Lunch?</td>
<td>State Math Score</td>
<td>St. Read Score</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>------------</td>
<td>-------------------------------</td>
<td>-----------------</td>
<td>---------------</td>
</tr>
</tbody>
</table>
**STUDENTS WITH ACCOMMODATIONS**

Please list students receiving accommodations in instruction and in testing and indicate the type(s) of accommodations.

<table>
<thead>
<tr>
<th>Student Name (Last Name, First Name)</th>
<th>Type(s) of Accommodations in Instruction and in Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Information regarding gender and ethnicity are optional.

**Ethnicity** - Please specify if the student is:
- a) Latino/Hispanic
- b) Caucasian
- c) African-American
- d) Asian/Pacific Islander
- e) Other (Please Specify)

**Disability Status**—Please include all disabilities listed in the student’s IEP.
- 1) Specific Learning Disabilities
- 2) Speech or Language Impairments
- 3) Mental Retardation
- 4) Emotional Disturbance
- 5) Deaf/Blindness (i.e. Students who are both deaf and blind)
- 6) Visual Impairments (Blind is included here)
7) Hearing Impairments (Deaf is included here)  
8) Orthopedic (Physical) Impairments  
9) Other Health Impairments  
10) Autism  
11) Traumatic Brain Injury  
12) Multiple Disabilities  
13) Developmental Delays  
14) Others (please specify)  

2 LEP Status—please specify current status:  
1) Limited English Proficient (LEP) – non-native English speakers  
2) Fluent English Proficient (FEP) – originally LEP, transitioned to FEP  
3) Initially Fluent in English (IFE) – native English speakers or bilingual students fluent in English when first tested.  

3 Receiving Free-Reduced Lunch—Please indicate YES or NO.  

4 State Math Score—Please provide most recent state math test score and please specify test used. Provide percentile score, if available, or other form (please specify).  

5 State Reading Score—Please provide most recent state reading test score and please specify test used. Provide percentile score, if available, or other form (please specify).  

6 Teacher Math Rating—Please rate the student with respect to state math standards for grade level, indicating whether the student is (if your state uses a different rubric, please note and explain levels below):  
   5 Highly Proficient  
   4 Proficient  
   3 Basic  
   2 Below Basic  
   1 Far Below Basic  

7 Teacher Reading Rating—Please rate the student with respect to state reading standards for grade level, indicating whether the student is (if your state uses a different rubric, please note and explain levels below):  
   5 Highly Proficient  
   4 Proficient  
   3 Basic  
   2 Below Basic  
   1 Far Below Basic  

Thank You
Appendix E: State Approved Accommodations

Appendix E: State-approved accommodations recommended for students with disabilities

<table>
<thead>
<tr>
<th>ID</th>
<th>Accommodations</th>
<th>ID</th>
<th>Accommodations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Extended Time</td>
<td>30</td>
<td>Copy of Lecture Notes</td>
</tr>
<tr>
<td>2</td>
<td>Shortened Assignment/Test</td>
<td>31</td>
<td>Work with ELL Teacher</td>
</tr>
<tr>
<td>3</td>
<td>Small Group</td>
<td>32</td>
<td>Translation of Vocabulary</td>
</tr>
<tr>
<td>4</td>
<td>Read Instructions and Test (if permitted)</td>
<td>33</td>
<td>Emphasis on Vocabulary</td>
</tr>
<tr>
<td>5</td>
<td>Frequent Breaks</td>
<td>34</td>
<td>Hands on Demonstration</td>
</tr>
<tr>
<td>6</td>
<td>Modified Grading</td>
<td>35</td>
<td>Fewer Vocabulary</td>
</tr>
<tr>
<td>7</td>
<td>Alternate Testing</td>
<td>36</td>
<td>Notes on Tests</td>
</tr>
<tr>
<td>8</td>
<td>Access to Resource Room</td>
<td>37</td>
<td>One-on-One Instruction</td>
</tr>
<tr>
<td>9</td>
<td>Repeat Instruction</td>
<td>38</td>
<td>Check for Understanding</td>
</tr>
<tr>
<td>10</td>
<td>Calculator Use on State Testing</td>
<td>39</td>
<td>Visuals/Gestures/Manipulation</td>
</tr>
<tr>
<td>11</td>
<td>Tutoring Assistance</td>
<td>40</td>
<td>Printed Directions</td>
</tr>
<tr>
<td>12</td>
<td>Clarify Directions</td>
<td>41</td>
<td>Graphic Organizers of the Text</td>
</tr>
<tr>
<td>13</td>
<td>Bi-lingual Explanation</td>
<td>42</td>
<td>Left Ear towards the Speaker</td>
</tr>
<tr>
<td>14</td>
<td>Preferential Seating</td>
<td>43</td>
<td>Paraphrasing</td>
</tr>
<tr>
<td>15</td>
<td>Alternate Assignment</td>
<td>44</td>
<td>Special Education</td>
</tr>
<tr>
<td>16</td>
<td>Math and Reading Lab below 4th Grade Work</td>
<td>45</td>
<td>SDAIE/SIOP/GLAD strategies generally include the following accommodations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9. Repeat Instructions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12. Clarify Directions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>33. Emphasis on Vocabulary</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>34. Hands on Demonstrations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>39. Visuals/Gestures/Manipulations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>41. Graphic Organizers of Text</td>
</tr>
<tr>
<td>17</td>
<td>Tested on Instructional Levels</td>
<td>46</td>
<td>Total Physical Response (TPR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>convey meaning by verbally associating an action/gesture with a picture (aka “realia”).</td>
</tr>
<tr>
<td>18</td>
<td>Study Sheets</td>
<td>47</td>
<td>Dictionary/Thesaurus</td>
</tr>
<tr>
<td>19</td>
<td>Teacher of Instruction for Test</td>
<td>48</td>
<td>Reader</td>
</tr>
<tr>
<td></td>
<td>Out-of-Level Tests</td>
<td></td>
<td>Guided Questions</td>
</tr>
<tr>
<td>---</td>
<td>--------------------------------------------------------</td>
<td>---</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>20</td>
<td>Out-of-Level Tests</td>
<td>49</td>
<td>Guided Questions</td>
</tr>
<tr>
<td>21</td>
<td>Use of an Assignment Notebook</td>
<td>50</td>
<td>Scribe</td>
</tr>
<tr>
<td>22</td>
<td>Work with other Hispanic students &amp; Hispanic Teacher</td>
<td>51</td>
<td>Scaffolded Teaching</td>
</tr>
<tr>
<td>23</td>
<td>Test correction or Retake</td>
<td>52</td>
<td>Allow Students to Speak in Native Language</td>
</tr>
<tr>
<td>24</td>
<td>Refocusing Attention to Task</td>
<td>53</td>
<td>Alternative Testing Site</td>
</tr>
<tr>
<td>25</td>
<td>Un-timed Testing</td>
<td>54</td>
<td>Large Print</td>
</tr>
<tr>
<td>26</td>
<td>Note-Taking Assistance</td>
<td>55</td>
<td>Shortened Lessons</td>
</tr>
<tr>
<td>27</td>
<td>Extra Instruction</td>
<td>56</td>
<td>Goal Setting/Objectives</td>
</tr>
<tr>
<td>28</td>
<td>Peer Tutoring (if necessary)</td>
<td>57</td>
<td>Slower Speech</td>
</tr>
<tr>
<td>29</td>
<td>Eliminate Choices on Multiple Choice Test</td>
<td>58</td>
<td>Oral Testing</td>
</tr>
</tbody>
</table>
Appendix F
Parent Information Letter
(English)

Dear Parent or Guardian,

Your child’s school is participating in a study, Validity of Accommodations for LEP Students with Disabilities in Math and English Proficiency Content Assessment. This study is examining the validity of test accommodations used for students with disabilities and English language learners in grades 4-12. The study is sponsored by the US Department of Education through the Oklahoma State Department of Education, coordinated by the Council of Chief State School Officers, and administered by Advance Research and Data Analyses Center (ARDAC).

Some students will take reading and writing tests, the English Language Development Assessment (two class periods), and some will take a mathematics test (approximately one class period). To examine the validity of accommodations, some students will be tested under accommodated conditions and some will not. Some students who have an Individualized Education Plan (IEP) will be tested without the accommodations specified in the IEP. The results of this study will be used for test development purposes. Your child’s scores will NOT be shared with your child’s teacher or principal. Test results will NOT become part of your child’s educational record.

Since it is not feasible to request prior parental permission for each child, this passive parental consent form is being provided. If for any reason you do NOT want your child to participate in the study, please complete the form below and return it to your child’s teacher within one week of receipt of this letter. If you do not return this form within one week, it will be assumed that you have no objection to your child’s participation in the study.

If you have any questions or concerns, please contact your child’s teacher, or telephone Fery Hejri (530)750-0625.

I DO NOT wish my child to participate in this study.

Print child’s name: _________________________________________
Print Child’s School name: ________________________________________
Print Child’s School address: ______________________________________

City: __________________ State: __________

Print name of parent or guardian: _____________________________

Signature of parent or guardian: _______________________________

Date: ___________
Appendix F: (English/Student)
Student Information Letter

Dear Student,

Your school is participating in a study, Validity of Accommodations for LEP Students with Disabilities in Math and English Proficiency Content Assessment. This study is examining the validity of test accommodations used for students with disabilities and English language learners in grades 4-12. The study is sponsored by the US Department of Education through the Oklahoma State Department of Education, coordinated by the Council of Chief State School Officers, and administered by Advance Research and Data Analyses Center (ARDAC).

Some students will take reading and writing tests, the English Language Development Assessment (two class periods), and some will take a mathematics test (approximately one class period). To examine the validity of accommodations, some students will be tested under accommodated conditions and some will not. Some students who have an Individualized Education Plan (IEP) will be tested without the accommodations specified in the IEP. The results of this study will be used for test development purposes. Your scores will NOT be shared with your teacher or principal. Test results will NOT become part of your educational record.

Since it is not feasible obtain prior permission from everyone, this passive student consent form is being provided. If for any reason you do NOT want to participate in the study, please complete the form below and return it to your teacher within one week of receipt of this letter. If you do not return this form within one week, it will be assumed that you have no objection to your participation in the study.

If you have any questions or concerns, please contact your teacher, or telephone Fery Hejri (530)750-0625 or Kathryn Lee at ARDAC, (310) 838-7927.

I DO NOT wish to participate in this study.

Print your name: _________________________________________
Print School name: _________________________________________
Print School address: _________________________________________

City: ________________State: ______________

Signature: _____________________________________

Date: ________________
Appendix F: (Spanish/Student)
Carta de Información para Estudiantes

Fecha:

Estimado Estudiante,

Su escuela está participando en un estudio, Validez de Adaptaciones para Estudiantes LEP con Incapacidades en Matemáticas y Valoración de Contenido de Habilidad en Inglés. Este estudio examina la validez de adaptaciones de exámenes que se usan para estudiantes con incapacidades y aprendices del idioma inglés en los grados escolares 4-12. El estudio está patrocinado por el Departamento Estadounidense de Educación a través del Departamento de Educación del Estado de Oklahoma, coordinado por el Concejo de Oficiales Escolares Estatales Principales, y administrado por Advance Research and Data Analyses Center [Centro Avanzado de Análisis de Investigación e Información] (ARDAC).

Algunos estudiantes tomarán pruebas de lectura y escritura, la Valoración de Desarrollo del Idioma Inglés [English Language Development Assessment] (dos períodos de clase), y algunos tomarán una prueba de matemáticas (aproximadamente un periodo de clase). Para examinar la validez de las adaptaciones, a algunos estudiantes se les darán pruebas bajo condiciones de adaptación y a algunos otros no. A algunos de los estudiantes que tienen un Plan Individualizado de Educación (IEP) se les darán pruebas sin la acomodación especificada en el IEP. Los resultados de este estudio servirán para propósitos de desarrollo de pruebas. Sus puntuaciones NO serán compartidas con su maestro o director. Los resultados de las pruebas no pasarán a formar parte de su archivo educativo.

Ya que no es factible obtener permiso previo de todos, se le está proporcionando este formulario pasivo de consentimiento estudiantil. Si por cualquier razón usted NO quiere participar en el estudio, por favor complete el formulario abajo y devuélvalo a su maestro a más tardar una semana de recibir esta carta. Si usted no devuelve este formulario a más tardar una semana, se supondrá que usted no tiene objeción a participar en el estudio.

Si tiene cualquier pregunta o inquietud, por favor comuníquese con su maestro, o llame por teléfono a Fery Hejri: (530)750-0625 o Kathryn Lee en ARDAC, (310) 838-7927. NO deseo participar en este estudio.

Escriba su nombre en letras de molde: ______________________________________

Firma: ___________________________________
Fecha: ________________
Appendix F: (Spanish/Parent)
Carta de Información para Padres

Fecha:

Estimado Padre o Tutor,

La escuela de su hijo está participando en un estudio, Validez de Adaptaciones para Estudiantes LEP con Incapacidades en Matemáticas y Valoración de Contenido de Habilidad en Inglés. Este estudio examina la validez de adaptaciones de exámenes que se usan para estudiantes con incapacidades y aprendices del idioma inglés en los grados escolares 4-12. El estudio está patrocinado por el Departamento Estadounidense de Educación a través del Departamento de Educación del Estado de Oklahoma, coordinado por el Concejo de Oficiales Escolares Estatales Principales, y administrado por Advance Research and Data Analyses Center [Centro Avanzado de Análisis de Investigación e Información] (ARDAC).

Algunos estudiantes tomarán pruebas de lectura y escritura, la Valoración de Desarrollo del Idioma Inglés [English Language Development Assessment] (dos periodos de clase), y algunos tomarán una prueba de matemáticas (aproximadamente un período de clase). Para examinar la validez de las adaptaciones, a algunos estudiantes se les darán pruebas bajo condiciones de adaptación y a algunos otros no. A algunos de los estudiantes que tienen un Plan Individualizado de Educación (IEP) se les darán pruebas sin la acomodación especificada en el IEP. Los resultados de este estudio servirán para propósitos de desarrollo de pruebas. Las puntuaciones de su hijo NO serán compartidas con el maestro o director de su hijo. Los resultados de las pruebas no pasarán a formar parte de su archivo educativo.

Ya que no es factible obtener permiso previo de los padres de todo niño, se le está proporcionando este formulario pasivo de consentimiento de padres. Si por cualquier razón usted NO quiere que su hijo participe en el estudio, por favor complete el formulario abajo y devuélvalo al maestro de su hijo a más tardar una semana de recibir esta carta. Si usted no devuelve este formulario a más tardar una semana, se supondrá que usted no tiene objeción a que su hijo participe en el estudio.

Si tiene cualquier pregunta o inquietud, por favor comuníquese con el maestro de su hijo, o llame por teléfono a Ferry Hejri (530) 750-0623 o Kathryn Lee en ARDAC, (310) 838-7927.

NO deseo que mi hijo participe en este estudio.

Escriba el nombre de su hijo en letras de molde: ______________________________

Escriba en letras de molde el nombre de padre o tutor: _______________________

Firma del padre o tutor: _____________________________________

Fecha: _______________
Appendix F: (Russian)

Информационное письмо для родителей

Уважаемый родитель или опекун!

Школа, в которой учится ваш ребенок, принимает участие в исследовании «Подтверждение наличия условий для оценки знаний по математике и английскому языку тех учащихся (с недостатками), которые имеют недостаточное знание английского языка». В рамках этого исследования изучается пригодность тестов для учащихся с недостатками и учащихся, изучающих английский язык в 4-12 классах. Спонсором исследования является министерство просвещения США. Оно финансирует исследование через управление образования штата Оклахома. Исследование координируется Советом главных руководителей школ штата, а само исследование проводится Центром перспективных исследований и анализа данных (ARDAC).

Некоторые учащиеся будут сдавать тесты по чтению и письму, пройдут оценку на развитие навыков владения английским языком (два классных периода), а некоторые будут сдавать тест по математике (примерно один классный период). Для подтверждения пригодности условий некоторые учащиеся будут сдавать тесты при принятых условиях, а некоторые нет. Некоторые учащиеся, имеющие индивидуальный план обучения (IEP), будут сдавать тесты без условий, указанных в IEP. Результаты этого исследования будут использоваться для совершенствования тестов. Оценки, полученные вашим ребенком, НЕ будут сообщаться преподавателю или директору школы, где учится ваш ребенок. Результаты тестов НЕ будут включены в документацию по обучению вашего ребенка.

Поскольку невозможно заранее запросить разрешение родителей каждого ребенка, мы предоставляем форму разрешения родителей по умолчанию. Если по какой-либо причине вы НЕ хотите, чтобы ваш ребенок участвовал в этом исследовании, заполните эту форму и передайте ее преподавателю вашего ребенка в течение одной недели после получения этого письма. Если вы не возвратите эту форму в течение одной недели, это будет означать, что вы не возражаете, чтобы ваш ребенок принял участие в исследовании.

Если у вас имеются какие-либо вопросы или соображения, просим связаться с преподавателем вашего ребенка или позвонить Фери Хейри по телефону 530-750-0625 или Кэтрин Ли по телефону 310:838-7927.

Я НЕ хочу, чтобы мой ребенок принимал участие в этом исследовании.

Написать печатными буквами имя и фамилию ребенка:
Написать печатными буквами название школы, где учится ваш ребенок:
__________________________________________
Написать печатными буквами адрес школы, где учится ваш ребенок:
__________________________________________
Город: _________________ Штат:_____________
Написать печатными буквами имя и фамилию родителя или опекуна ребенка:
__________________________________________
Подпись родителя или опекуна: _________________________________
____
Дата: ________________
Appendix G:

TEST ADMINISTRATOR TRAINING AGENDA
FOR ACCOMMODATIONS STUDY

I. General Instructions for the ELDA and Math Tests

Security:
You are responsible for the security of these confidential test materials. Measurement Incorporated and ARDAC request that you keep all test materials in an extremely safe and secure environment. Do not photocopy them or allow anyone access to them other than yourself and the school contact person.

Preparing for the test:
It is very important to familiarize yourself with the scripted directions you will be reading to the students before the day of the test as well as before your teleconference training. You may also want to look over the Student Background Questionnaire so that you are prepared to answer students’ questions as they fill them out. Thorough preparation will prevent delays and will minimize the amount of time students are away from class.

Review the accommodations your students require and make sure that you are prepared accordingly. For instance, you will need to set up a separate room or area for students who require a small group or alternate location testing setting. Also, be sure to enlist as many people to help you administer the test as you may need, such as someone to help with timing an alternate group of students or reading directions. Prepare the conditions that your accommodated students will require in accordance with the test’s parameters, and be careful not to offer any assistance that the test does not permit. Test directions may not be translated into a language other than English.

Before the test:
Verify the materials received against the packing slip. Make sure you have the appropriate materials in sufficient quantities for testing students in your classroom. For each test you will need:

- Test booklets
- Sharpened no.2 pencils and erasers
- Stopwatch or standard watch

During the test:
Be careful to preserve properly accommodated conditions without compromising the security of the test. Divide classroom between bilingual and non-bilingual test takers so that you can manage students who have completed the test. Designate an area in your classroom where students who have completed the test can read quietly or work
independently on other assignments, or permit these students to read or work at their own desks as long as they are quiet.

After the test:
It is imperative to confirm that you have the exact same number of booklets after testing as the initial number of booklets sent to you. MI must confirm that they have received all test materials, used and unused, before reimbursements are disbursed to schools, and will conduct a security check-in upon receipt of the test materials.

II. The ELDA Test

The test itself:
ELDA tests two subjects: reading, and writing. A chart of the estimated maximum time necessary for the test is below.

<table>
<thead>
<tr>
<th>ELDA class</th>
<th>Student Background Questionnaire</th>
<th>Testing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular student</td>
<td>Approximately 10 minutes; allow time for questions</td>
<td>2 hours</td>
</tr>
<tr>
<td>Students with accommodations</td>
<td>Approx. 10 minutes; allow time for questions</td>
<td>3 hours</td>
</tr>
</tbody>
</table>

Sargent: You have 4 students who require extended time. 2 of these students require a small group seating arrangement, and 2 require preferential seating. Please prepare proper seating and grouping arrangements and allow enough time to accommodate these students.

LiRung: You have 2 students who require extra time and preferential seating. Please prepare proper seating and grouping arrangements and allow enough time to accommodate these students.

III. The Math Test

The test itself:
The math test assesses students’ performance in grade-level math. A chart of the estimated maximum time necessary for the test is below.

<table>
<thead>
<tr>
<th>Math class</th>
<th>Student Background Questionnaire</th>
<th>Testing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math test 1 or 3 (English)</td>
<td>Approximately 10 minutes; allow time for questions</td>
<td>35 minutes</td>
</tr>
<tr>
<td>Bilingual Version</td>
<td>Approximately 50 minutes</td>
<td></td>
</tr>
<tr>
<td><strong>Spanish &amp; English</strong></td>
<td>10 minutes; allow time for questions</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>SD with Accommodations</strong></td>
<td>Approximately 10 minutes; allow time for questions</td>
<td>50 minutes (if they have extra time)</td>
</tr>
</tbody>
</table>

Warren: You have 10 students who require extended time. 6 of these students require a small group, 2 of these students require preferential seating, and 4 of these students require a calculator. You have 4 students who will be taking the Bilingual Side-by-Side English-Spanish version of the math test and will require extended time. Please prepare proper seating and grouping arrangements, provide calculators or remind students to bring their own, and allow enough time to accommodate these students.
Appendix H:

Validity of Accommodations for LEP Students with Disabilities in Math and English Proficiency Content Assessment

The Oklahoma State Department of Education
Council of Chief State School Officers
Advance Research & Data Analyses Center (ARDAC)

High School Participation Information Sheet

School Responsibilities:
- The study needs access to a maximum of 2 high schools with the largest concentration of LEP students.
- Each school will provide access to a maximum of two classes in grade 11 and two classes in grade 12. [Figure.doc]
- All students in intact classroom will be tested including students with limited English proficiency, students with disabilities, and regular students.
- School will identify a School Contact Person who will coordinate school testing.
- Will complete School Questionnaire (by School Principal). [School_Questionnaire.pdf]

School Contact Person’s Responsibilities:
- Will establish contact with ARDAC to receive instruction.
- Will identify four classrooms to be tested and determine testing date and time for each of the four participating classrooms.
- Will receive a Class Roster and distribute to each of the four participating classrooms then will fax and/or mail them to ARDAC two weeks prior to test administration. All information provided will be used for the purposes of this project only and will be kept strictly confidential. [Class_Roster4- 5-7-8-11-12 new.pdf]
- Will identify a Test Administrator who will then be trained by ARDAC research staff.
- Will work with the Test Administrator in implementing student IEP accommodations.
- Will provide school contact information (address and telephone number) to ARDAC for shipment of all test materials.
- Will receive a $250 gift card (from Office Depot, Staples, etc.) or cash (if acceptable based on state policy).
Test Administrator’s Responsibilities:

- Will participate in a training session conducted by ARDAC. Training will be done either on site by ARDAC research staff or by telephone. Will receive a packet of training materials one week prior to the training session.
- Will receive instructions from ARDAC for providing accommodations to students with disabilities and LEP students.
- Will receive all test materials from School Contact Person.
- Will administer ELDA Test and Student Background Questionnaire to students.
  - ELDA Test will be shipped by MI to School Contact Person.
  - Student Questionnaire [Student_Questionnaire.pdf]
- Will collect all test materials from each participating classroom and mail test materials to Measurement Incorporated (MI) using the pre-paid envelopes provided:
  Measurement Incorporated
  423 Morris Street
  Durham, NC 27701
- Will receive $375 ($150 per day for two days plus $75 for half a day of training).

Teacher’s Responsibilities:

- Will provide Test Administrator access to his/her class.
- Distribute Parent Assent and Student Assent forms for all students that will be participating (if acceptable based on state policy). Collect ant returned Assent forms; make alternative arrangements for any students who will not be participating in the testing.
  - Student Assent English version [Student-Assent-English-N.pdf]
  - Parent and Student Assent Spanish version [Parent-Assent-Span-N.pdf]
  - Parent Assent Russian version [Parent-Assent-Russian-N.pdf]
- Will assist Test Administrator in administering the test materials with or without accommodations.
- Will complete Class Roster Questionnaire.
- Will complete Teacher Questionnaire [Teacher_Questionnaire.pdf]
- Will receive a $120 gift card (from Office Depot, Staples, etc.).

Contact Information:
Fery Hejri at info@ardacresearch.com or
Kathryn Lee at kathryn@ardacresearch.com or
Advance Research & Data Analysis Center
2062 John Jones Rd, Suite 230
Appendix I:

Validity of Accommodations for LEP Students with Disabilities in Math and English Proficiency Content Assessment

The Oklahoma State Department of Education
Council of Chief State School Officers
Advance Research & Data Analyses Center (ARDAC)

Middle School Participation Information Sheet

School Responsibilities:
- The study needs access to a maximum of two middle schools with the largest concentration of LEP students.
- Each school will provide access to a maximum of two classes in grade 7 and two classes in grade 8. [Figure.doc]
- All students in intact classroom will be tested including students with limited English proficiency, students with disabilities and regular students.
- School will identify a School Contact Person who will coordinate school testing.
- Will complete School Questionnaire(by School Principal). [School_Questionnaire.pdf]

School Contact Person’s Responsibilities:
- Will establish contact with ARDAC to receive instruction.
- Will identify four classrooms to be tested and determine testing date and time for each of the four participating classrooms.
- Will receive a Class Roster and distribute to each of the four participating classrooms then will fax and/or mail them to ARDAC two weeks prior to test administration. All information provided will be used for the purposes of this project only and will be kept strictly confidential. [Class_Roster4- 5-7-8-11-12 new.pdf]
- Will identify a Test Administrator who will then be trained by ARDAC research staff.
- Will work with the Test Administrator in implementing student IEP accommodations.
- Will provide school contact information (address and telephone number) to ARDAC for shipment of all test materials.
- Will receive a $250 gift card (from Office Depot, Staples, etc.) or cash (if acceptable based on state policy).

Test Administrator’s Responsibilities:
• Will participate in a training session conducted by ARDAC. Training will be
done either on site by ARDAC research staff or by telephone. Will receive a
packet of training materials a week prior to the training session.

• Will receive instructions from ARDAC for providing accommodations to
students with disabilities and LEP students.
• Will receive all test materials from School Contact Person.
• Will administer ELDA Test\(^2\), Math Test, and Student Background Questionnaire
to students.
  o ELDA Test will be shipped by MI to School Contact Person.
  o Math Test will be shipped by ARDAC to School Contact Person.
  o Student Questionnaire [Student_Questionnaire.pdf]
• Will collect all test materials from each classroom and mail math test materials to
ARDAC using the pre-paid envelopes provided (test materials for each
classroom should be placed in separate envelopes). ELDA Test booklet only
should be mailed to:
Measurement Incorporated
423 Morris Street
Durham, NC 27701
• Will receive $375 ($150 per day for two days plus $75 for half a day of training).

Teacher’s Responsibilities:
• Will provide Test Administrator access to his/her class.
• Distribute Parent Assent and Student Assent forms for all students that will be
participating (if acceptable based on state policy). Collect and returned Assent
forms; make alternative arrangements for any students who will not be
participating in the testing.
  o Student Assent English version [Student-Assent-English-N.pdf]
  o Parent and Student Assent Spanish Version [Parent-Assent-Span-N.pdf]
  o Parent Assent Russian version [Parent-Assent-Russian-N.pdf]
• Will assist Test Administrator in administering the test materials with or without
accommodations.
• Will Complete Class Roster Questionnaire.
• Will complete Teacher Questionnaire. [Teacher_Questionnaire.pdf]
• Will receive $120 in gift card (from Office Depot, Staples, etc.).

Contact Information:

\(^2\) MI will prepare ELDA Test materials
Fery Hejri at info@ardacresearch.com or
Kathryn Lee at kathryn@ardacresearch.com
Advance Research & Data Analysis Center
2062 John Jones Rd, Suite 230
Davis, CA 95616
Telephone: (530) 750-0625
Fax: (530) 750-0623
Appendix J:

Validity of Accommodations for LEP Students with Disabilities in Math and English Proficiency Content Assessment

The Oklahoma State Department of Education
Council of Chief State School Officers
Advance Research & Data Analyses Center (ARDAC)

Elementary School Participation Information Sheet

School Responsibilities:
- The study needs access to a maximum of two elementary schools with the largest concentration of LEP students.
- Each school will provide access to a maximum of two classes in grade 4 and two classes in grade 5. [Figure.doc]
- All students in intact classroom will be tested including students with limited English proficiency, students with disabilities and regular students.
- School will identify a School Contact Person who will coordinate school testing.
- Will complete School Questionnaire (by School Principal). [School_Questionnaire.pdf]

School Contact Person’s Responsibilities:
- Will establish contact with ARDAC to receive instruction.
- Will identify four classrooms to be tested and determine testing date and time for each of the four participating classrooms.
- Will receive a Class Roster and distribute to each of the four participating classrooms then will fax and/or mail them to ARDAC two weeks prior to test administration. All information provided will be used for the purposes of this project only and will be kept strictly confidential. [Class_Roster4- 5-7-8-11-12 new.pdf]
- Will identify a Test Administrator who will then be trained by ARDAC research staff.
- Will work with the Test Administrator in implementing student IEP accommodations.
- Will provide school contact information (address and telephone number) to ARDAC for shipment of all test materials.
- Will receive a $250 gift card (from Office Depot, Staples, etc.) or cash (if acceptable based on state policy).

Test Administrator’s Responsibilities:
- Will participate in a training session conducted by ARDAC. Training will be done either on site by ARDAC research staff or by telephone. Will receive a packet of training materials one week prior to the training session.
- Will receive instructions from ARDAC for providing accommodations to students with disabilities and LEP students.
- Will receive all test materials from School Contact Person.
- Will administer ELDA Test, Math Test, and Student Background Questionnaire to students.
  - ELDA Test will be shipped by MI to School Contact Person.
  - Math Test will be shipped by ARDAC to School Contact Person.
  - Student Questionnaire [Student_Questionnaire_4-5_7-8-11-12.pdf]
- Will collect all test materials from each participating classroom and mail math test materials to ARDAC using the pre-paid envelopes provided (test materials for each classroom should be placed in separate envelopes). ELDA Test booklet only should be mailed to:
  Measurement Incorporated
  423 Morris Street
  Durham, NC 27701
- Will receive $375 ($150 per day for two days plus $75 for half a day of training).

**Teacher’s Responsibilities:**
- Will provide Test Administrator access to his/her class.
- Distribute Parent Assent and Student Assent forms for all students that will be participating (if acceptable based on state policy). Collect ant returned Assent forms; make alternative arrangements for any students who will not be participating in the testing.
  - Student Assent English version [Student-Assent-English-N.pdf]
  - Parent and Student Assent Spanish version [Parent-Assent-Span-N.pdf]
  - Parent Assent Russian version [Parent-Assent-Russian-N.pdf]
- Will assist Test Administrator in administering the test materials with or without accommodations.
- Will complete Class Roster Questionnaire.
- Will complete Teacher Questionnaire. [Teacher_Questionnaire.pdf]
- Will receive a $120 gift card (from Office Depot, Staples, etc.).

**Contact Information:**

---

3 MI will prepare ELDA Test materials
Fery Hejri at info@ardacresearch.com or
Kathryn Lee at kathryn@ardacresearch.com
Advance Research & Data Analysis Center
2062 John Jones Rd, Suite 230
Davis, CA 95616
Telephone: (530) 750-0625
Fax: (530) 750-0623
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>3</td>
</tr>
<tr>
<td>General Test Information</td>
<td>3</td>
</tr>
<tr>
<td>Responsibility of Test Administrator</td>
<td>3</td>
</tr>
<tr>
<td>Before Test Administration</td>
<td>3</td>
</tr>
<tr>
<td>During Test Administration</td>
<td>4</td>
</tr>
<tr>
<td>After Test Administration</td>
<td>5</td>
</tr>
<tr>
<td>Students with Accommodations</td>
<td>5</td>
</tr>
<tr>
<td>Test Administrator Checklist</td>
<td>6</td>
</tr>
<tr>
<td>Grade 5 Math Directions</td>
<td>7</td>
</tr>
<tr>
<td>Grade 8 Math Directions</td>
<td>10</td>
</tr>
</tbody>
</table>
INTRODUCTION

This Test Administration Manual contains general information about the test and information about administering the test to grades 5 and 8. Before administering the test please carefully review the information in this manual to familiarize yourself with the test administration procedures.

GENERAL TEST INFORMATION

This is a 30-item math test in two different forms: (1) the original English version, and (2) the English-Spanish bilingual version. Form 1 (the English version) should take approximately 35 minutes to complete. Students receiving Form 2 (the Bilingual version) have the option of reading items and responding either in English or Spanish. Therefore, they need additional time to respond to this version of the test. For this version (Form 2), we allocated 50 minutes. There will be no breaks during the test.

Advance Research and Data Analyses Center (ARDAC) will send all test materials to your school prior to the test administration day. It is your most important responsibility as a test administrator to keep all materials in a secure and safe place designated by your School Contact Person. No student should have access to any of the test materials before and after the test.

RESPONSIBILITIES OF THE TEST ADMINISTRATOR

BEFORE TEST ADMINISTRATION

The following materials will be shipped to your School Contact Person and it is the Test Administrator's responsibility to retrieve these materials:

- Math test booklets and test instruction manual
- Pre-paid envelopes to return Math test booklets and Questionnaires
- Class Roster – to be completed by the Classroom Teacher and must be returned to ARDAC **two weeks prior** to the test administration.
- School Questionnaire – to be completed by the School Principal
- Teacher Questionnaire – to be completed by the Classroom Teacher
- Student Questionnaire – to be completed by each individual participating student
Once all test materials have been accounted for, the Classroom Roster, School Questionnaire, Teacher Questionnaire, and Student Questionnaire should be administered to the appropriate persons. The Class Roster must be completed by the classroom teacher and returned to ARDAC two weeks prior to the test administration.

On the day of the test administration, please arrive early to ensure that you have all the following materials available:

- Math test booklets
- Sharpened No. 2 pencils with erasers
- A stopwatch or standard watch

Check the assessment room to see that all visual clues and aids are removed or covered and remain hidden throughout the test administration. Once the students arrive, make sure that their desks are cleared of any materials that are unrelated to the test.

**DURING TEST ADMINISTRATION**

Inform students that once they receive their test booklet there should be absolutely no talking until the test is completed and everyone has turned in their test booklet.

Once the test administration has begun, walk around the classroom to ensure that all students are marking the test booklet correctly (circling and filling in the answers in the test booklet) and that they are not looking at other tests. All responses must be marked in the test booklet or they will not be scored. Encourage students to answer all questions and not leave any blank.

If a student has a question during the test, please ask him/her to quietly raise their hand. You may answer all questions related to completing the test, but do not provide assistance on specific test questions.

Students are allowed to go to the restroom during the test; however, if more than one student requests to go to the restroom only one student should leave the room at any one time. As a test administrator you must remain in the classroom at all times and will only be allowed to leave once all test materials have been collected and accounted for.
During the test there may be students who finish the test before others. Please request that they remain quiet while others complete their test. You may allow those students who finish early to read a book or do other school work just as long as it is unrelated to the test material content. Remember to collect all test booklets from those students who finish early before they are allowed to do other work.

AFTER TEST ADMINISTRATION

Once time is called, have all students close their test booklet and continue to remain quiet while you collect their booklets. Before students are allowed to leave the classroom, ensure that all test booklets have been returned. All test materials (used and unused) should then be placed into the pre-paid envelope and returned to ARDAC. The School Questionnaire, Teacher Questionnaire, and Student Questionnaire must also be returned to ARDAC in the pre-paid envelope provided.

STUDENTS WITH ACCOMMODATIONS

An accommodation is a change in the test administration or testing environment. It does not change the content of the assessment or what the assessment measures. Appropriate accommodations should provide students the opportunity to demonstrate their knowledge and skills. Accommodations that are allowable for students with disabilities during assessment are those that are specified in a student’s IEP or 504 plan. The accommodation for assessment must also be consistent with the current instructional and classroom accommodations.

Any accommodation provided to a student must be specified prior to the test administration and must be documented in the student’s IEP or 504 plan. It must also be recorded in the Class Rosters.

Test administrators will be informed in the training session about the type and administration of accommodations.
# TEST ADMINISTRATOR CHECKLIST

## BEFORE TESTING

- Participate in the training session a week prior to test administration (by this time the classroom teacher has completed the Class Roster).
- Arrange for sufficient personnel to supervise the testing.
- Review test administration directions.
- Prepare testing location.
- Administer Student Questionnaire to participating students.

### Materials:
- Math test booklets
- Sharpened No. 2 pencils with erasers
- A stopwatch or standard watch

## DURING TESTING

- Administer test.
- Use appropriate accommodations (this will be discussed in the test administrator training session).

## AFTER TESTING

- Collect all test booklets from each participating classroom.
- Count the test booklets to ensure that all have been returned.
- Return all test materials (used and unused) from each participating classroom to ARDAC in the pre-paid envelope provided.
- Complete and return the School Questionnaire, Teacher Questionnaire, and Student Questionnaire to ARDAC in the pre-paid envelope provided.
- Test Administrator will be paid $375 ($150 per day for two days of test administration and $75 for half day training).
GRADE 5 MATH DIRECTIONS

Begin administering the math test once every student is properly seated at their desks. Make sure that all students have a No. 2 pencil with an eraser. The test should take approximately one class period (35 minutes) to complete.

Please try to follow the directions as closely as possible. All scripted directions are in **bold print** and should be read verbatim.

**Math Test Scripted Directions**

You are about to take a 30-item Math test either in English or in Bilingual form (both in English and Spanish). The English version of the test will take approximately 35 minutes to complete the test. Those of you who are assigned to the Bilingual version of the test will have the option of reading and responding to the math items either in English or in Spanish. For this version you have 50 minutes to complete the test. As soon as the test is handed out to you please remain quite throughout the test. If you have any questions please raise your hand quietly and I will come around to help you. Do not open the test booklet until I tell you to do so. Pass out the Math test booklets. **Does everyone have a No. 2 pencil with an eraser?** Hand out extra pencils and erasers if necessary.

On the very first page of your test booklet please write your name, your teacher's name, your school name, your school district, and your state in the **blank spaces provided**. Demonstrate by using a sample test booklet or borrow from a student. Wait for every student to complete this part and provide help if necessary.

**You may now open your test booklet to the second page.** Make sure all students have the correct page showing. **Please read through each math question carefully.** Answer all questions as well as you can and do not spend too much time on any one question. If you skip a question, you may go back and answer it. When you answer the questions you MUST mark the test booklet otherwise your
answer will not be scored. For example, circle the letter with the best answer for multiple choice questions or write your answer on the blank line provided in your test booklet. If you change your answer, please erase your answer completely and then make your corrections.

Let’s do a couple of examples before you begin the test:
Sample Question 1. This question requires you to circle the correct answer. Read the sample to yourself as I read it aloud.

Martha planted 32 seeds. She put 8 seeds in each row. How many rows did she plant?
A) 32 + 8
B) 32 – 8
C) 32 x 8
D) 32 ÷ 8

Pause for replies. **Answer D is the correct answer.** In your test booklet you should circle letter D. You may circle just the letter D or you can circle the entire answer.

**Turn to page 3.** Make sure all students have the correct page showing. **Let’s do Sample Question 2.** This question requires you to write your answer on the blank line provided in your test booklet. Read the sample to yourself as I read it aloud.

Add 32 and 14
Answer: __________

Write the correct answer on the line provided. **Pause for replies. The correct answer is 46.** In your test booklet you should write 46 on the line provided. **Are there any questions?** You may answer questions about how to mark the test booklet, but do not assist students with the mathematics of the question.
In a few moments you will begin the math test. Remember, talking is not allowed during the test. If you have any questions please raise your hand and I will come around to help you. Also, remember to read through each question carefully and answer all questions as well as you can. Remember to circle your answers or write on the blank line provided in your test booklet.

If you finish before time is called please raise your hand and I will come and collect your test booklet. You may then read a book or do other school work. But please continue to remain quiet because other students around you may still be taking the test.

Are there any questions? Answer any additional questions related to completing the test. Do not provide assistance on specific test questions. You may now turn your test booklet to page 4 and begin the test. Please keep track of time.

GRADE 8 MATH DIRECTIONS

Begin administering the math test once every student is properly seated at their desks. Make sure that all students have a No. 2 pencil with an eraser. The test should take approximately one class period (35 minutes) to complete.

Please try to follow the directions as closely as possible. All scripted directions are in bold print and should be read verbatim.

Math Test Scripted Directions

You are about to take a 30-item Math test either in English or in Bilingual forms (both in English and Spanish). The English version of the test will take approximately 35 minutes to complete the test. Those of you who are assigned to the Bilingual version of the test will have the option of reading and responding to the math items either in English or in Spanish. For this version you have 50 minutes to complete the test. If you have any questions please raise your hand quietly and I will come around to help you. Do not open the test booklet until I tell you to do so. Pass out the Math test booklets. Does everyone have a No. 2 pencil with an eraser? Hand out extra pencils and erasers if necessary.
On the very first page of your test booklet please write your name, your teacher’s name, your school name, your school district, and your state in the blank spaces provided. Demonstrate by using a sample test booklet or borrow from a student. Wait for every student to complete this part and provide help if necessary.

You may now open your test booklet to the second page. Make sure all students have the correct page showing. Please read through each math question carefully. Answer all questions as well as you can and do not spend too much time on any one question. If you skip a question, you may go back and answer it. When you answer the questions you MUST mark the test booklet otherwise your answer will not be scored. For example, circle the letter with the best answer for multiple choice questions or write your answer on the blank line provided in your test booklet. If you change your answer, please erase your answer completely and then make your corrections.

Let’s do a couple of examples before you begin the test:
Sample Question 1. This question requires you to circle the correct answer. Read the sample to yourself as I read it aloud.

The average weight of 50 prize-winning tomatoes is 2.36 pounds. What is the combined weight, in pounds, of these 50 tomatoes?

A) 0.0472  
B) 11.8  
C) 52.36  
D) 59  
E) 118

Pause for replies. Answer E is the correct answer. In your test booklet you should circle letter E. You may circle just the letter E or you can circle the entire answer.
Turn to page 3. Make sure all students have the correct page showing. Let’s do Sample Question 2. This question requires you to write your answer on the blank line provided in your test booklet.

Add 32 and 14.
Answer: ________

Write the correct answer on the line provided. Pause for replies. The correct answer is 46. In your test booklet you should write 46 on the line provided. Are there any questions? You may answer questions about how to mark the test booklet, but do not assist students with the mathematics of the question.

In a few moments you will begin the math test. Remember, talking is not allowed during the test. If you have any questions please raise your hand and I will come around to help you. Also, remember to read through each question carefully and answer all questions as well as you can. Remember to circle your answers or write on the blank line provided in your test booklet.

If you finish before time is called please raise your hand and I will come and collect your test booklet. You may then read a book or do other school work. But please continue to remain quiet because other students around you may still be taking the test.

Are there any questions? Answer any additional questions related to completing the test. Do not provide assistance on specific test questions. You may now turn your test booklet to page 4 and begin the test. Please keep track of time.

Thank You
Appendix L:
Math Tests Grade 5


http://www.ardacresearch.com/Grade 5 LargePrint Form B-2007.pdf

http://www.ardacresearch.com/Large_Print_Grade_5_LanguageModified_Form_B-2007.pdf

http://www.ardacresearch.com/Large_Print_Grade_5_LanguageModified_Form_A-2007.pdf

http://www.ardacresearch.com/MATH Grade5 FormA_Spanish and English.pdf