



THE GEORGE WASHINGTON UNIVERSITY
CENTER FOR EQUITY AND
EXCELLENCE IN EDUCATION

Best Practices in State Assessment Policies for Accommodating English Language Learners: A Delphi Study

Barbara D. Acosta, Charlene Rivera, and Lynn Shafer Willner

October 2008

GW-CEEE
The George Washington University
Center for Equity and Excellence in Education
1555 Wilson Boulevard, Suite 515
Arlington, VA 22209
Tel: 703.528.3588
Fax: 703.528.5973
<http://ceee.gwu.edu>

**The George Washington University
Center for Equity and Excellence in Education**

The mission of The George Washington University Center for Equity and Excellence in Education (GW-CEEE) is to advance education reform so all students achieve to high standards. GW-CEEE conducts policy and applied research, designs and implements program evaluations, and provides professional development and technical assistance. The Center's clients include state education agencies, school districts, schools, foundations, and federal agencies.

For over a decade, GW-CEEE has conducted research on the inclusion and accommodation of English language learners (ELLs) in high stakes testing, including periodic reviews of state assessment policies for ELLs. GW-CEEE is currently conducting an applied research project with support from the Bill & Melinda Gates Foundation to study the academic language demands of ELLs in middle and high schools. GW-CEEE also provides technical assistance and conducts professional development for clients in states, districts, and schools. With funding from the U.S. Department of Education, GW-CEEE operates the Mid-Atlantic Comprehensive Center (MACC) and the Mid-Atlantic Equity Assistance Center (MEAC), both of which provide technical assistance to states in the Mid-Atlantic region.

Acknowledgements

This publication was supported by the U.S. Department of Education's (ED) LEP Partnership, a partnership of the states with the Deputy Secretary's Office and the Office of English Language acquisition at ED, the National Clearinghouse for English Language Acquisition, the Council of Chief State School Officers, the National Council of La Raza, and the Mexican American Legal Defense Fund.

As part of Partnership meetings, external assessment experts are made available to states to discuss the quality of the assessment process for English language learners (ELLs) and to examine how state assessments can be made more accessible to ELLs. This particular project grew out of discussions with state education agency staff including assessment, Title I, and III directors who had concerns about selecting and using accommodations for ELLs.

Identifying accommodations for ELLs that meet requirements of both effectiveness and validity is a challenging task. In the absence of a strong research base, ED charged The George Washington University Center for Equity and Excellence in Education (GW-CEEE) to develop a *Guide for Refining State Assessment Policies for Accommodating ELLs*. To prepare the foundation for the Guide, GW-CEEE conducted two studies, a Descriptive study of state policies and a Delphi study of practice.

We are grateful for the collaboration among ED offices that believed it was important to fund work in this area. In particular we would like to thank Ray Simon, Deputy Secretary; Kathryn Doherty, Special Assistant Office of the Deputy Secretary; Richard Smith, Acting Director for the Office of English Language Acquisition (OELA); Jenelle Leonard, Director, School Support and Technology Programs; and Fran Walter, Group Leader, Technical Assistance Programs.

In particular, the GW-CEEE research team is appreciative to staff from the 51 SEAs who cooperated in providing state policies and verifying the accuracy of state policy data. Their feedback on drafts of the Descriptive Study was invaluable.

We are indebted to the members of the expert panel. With their help, the GW-CEEE research team sorted and ranked ELL-responsive accommodations and mapped them to ELP levels. They provided valuable feedback on drafts of the two study reports. We would also like to recognize Arthur Dole, Professor Emeritus of Educational Psychology at the University of Pennsylvania; Dave Schwandt, Professor of Human and Organizational Studies at The George Washington University, and Charles Stansfield, President of Second Language Testing, Inc for their inspiration and advice.

Special recognition goes to GW-CEEE staff who provided important support throughout the project. Diane Staehr Fenner coordinated the development of an English language proficiency rubric for the Delphi Study, provided assistance during the data collection process, and gave feedback on the first draft of the Descriptive Study. Roshaun Tyson provided design and layout expertise. Kristina Anstrom, Abhee Brahmalkar, Janet Brown, Marilyn Muirhead, James Ruff, and Tracy Runfola helped pilot the Delphi questionnaires and provided feedback on drafts of the two studies and the Guide; Jennifer Lynch and Gamaliel Gabriento provided administrative support; and Charlotte Blane provided expert editorial assistance.

Expert Panel

Jamal Abedi, Ph.D.
Professor
School of Education
University of California, Davis

Linda Carstens, Ph.D.
Director
Stanford LEADS Network
Stanford University

H. Gary Cook, Ph.D.
Researcher
Wisconsin Center for Education Research
University of Wisconsin

Richard Duran, Ph.D.
Professor
University of California, Santa Barbara

Margo Gottlieb, Ph.D.
Director of Assessment and Evaluation
Illinois Resource Center

Robin Lisboa, M.S.
Administrator
Division of English Language Learning
Illinois State Board of Education

Lorena Llosa, Ph.D.
Assistant Professor
New York University

Carlos Martinez, M.A.
Assistant Secretary of Education
New Mexico Public
Education Department

Teddi Predaris, M.S., M.A.
Director
Office of ESOL Services
Fairfax County Public Schools

Maria Pennock-Roman, Ph.D.
Research Consultant
MPR Psychometric & Statistical Research
Consulting

Raquel Sinai, M.Ed.
Coordinator
Bilingual/ESL Education
New Jersey Department of Education

Table of Contents

Acknowledgments.....	iii
Expert Panel.....	iv
Table of Contents.....	v
List of Tables.....	vi
List of Figures.....	vi
Executive Summary.....	vii
Introduction.....	1
Review of the Literature.....	2
ELL Responsive Taxonomy.....	3
Effectiveness of Accommodations for ELLs.....	3
Mapping Accommodations to a Heterogeneous Population.....	5
Review of Assessment Policies for ELLs.....	6
Method.....	7
Limitations of the Study.....	9
Results of the Delphi Process.....	10
ELL-Responsive Accommodations.....	10
Accommodations Mapped to ELP Levels and Student Profiles.....	10
Beginning ELP Level.....	14
Intermediate ELP Level.....	16
Advanced ELP Level.....	16
Prior Experience Requirement.....	17
Accommodations Removed During the Screening Process.....	17
Accommodations not considered ELL-responsive.....	17
Accommodations that might threaten validity.....	19
Discussion.....	20
ELL-Responsiveness.....	20
Mapping Accommodations to ELP and Literacy Levels.....	21
Avoiding Threats to Validity.....	21
Test Administration Practices.....	22
Recommendations.....	23
Screen Accommodations for ELL-Responsiveness.....	23
Specify Accommodations to Used for Content Assessments.....	24
Standardize and Clearly Describe Accommodations.....	24
Distinguish Between Test Administration Practices and Accommodations.....	25
Offer Accommodations for ELLs at Each English Language Proficiency Level.....	25
Offer Accommodations for ELLs with Different Levels of Literacy in English and the Native Language.....	26
Future Research.....	26
References.....	27

List of Tables

Table 1. Direct and Indirect Linguistic Support Accommodations for English and Native Language.....	4
Table 2. Summary of Delphi Procedures	9
Table 3. ELL-Responsive Accommodations	11
Table 4. GW-CEEE English Language Proficiency Rubric	12
Table 5. Student Profiles Used for Mapping Accommodations in the Native Language.....	13
Table 6. Mapping of Accommodations in English to ELP levels.....	14
Table 7. Mapping of Native Language Accommodations to ELP levels and Student Profiles	15
Table 8. Accommodations Not Considered ELL-Responsive.....	18
Table 9. Test Administration Practices for Specific Accommodations	23

List of Figures

Figure 1. Conceptual Framework for Mapping Accommodations to ELP levels.....	13
--	----

Executive Summary

Charged by the U.S. Department of Education, The George Washington University Center for Equity and Excellence in Education (GW-CEEE) developed a Guide for state education agencies (SEAs) to use to improve state assessment policies for accommodating English language learners (ELLs). As a foundation for the Guide, GW-CEEE designed two studies, the Descriptive Study and the Best Practices Study. For the Descriptive Study GW-CEEE reviewed state assessment policies and examined the number and types of accommodations specified for ELLs (Shafer Willner, Rivera, & Acosta, 2008). The Best Practices Study involved the application of a Delphi technique to obtain consensus from an expert panel about which accommodations identified in the Descriptive Study were ELL-responsive. Members of the panel, which included experts knowledgeable about research, policy and practice in the areas of assessment, psychometrics, language testing, second language acquisition, and instruction of ELLs, relied on professional judgment to vet a list of ELL-responsive accommodations and then mapped these accommodations to English language proficiency (ELP) levels and to selected student background variables.

Summary of Findings from the Best Practices Study

ELL-Responsive Accommodations

The expert panel concluded that 29 of the accommodations identified in the Descriptive Study (taken from SEA policies) met the operational definition of ELL-responsive, i.e., accommodations likely to reduce *construct-irrelevant variance* due to language. The Best Practices Study contains the list of the accommodations organized according to an ELL-responsive taxonomy. In contrast with the traditional taxonomy used for accommodations for students with disabilities (presentation, response, timing/scheduling and setting), the ELL-responsive taxonomy draws attention to the *linguistic* needs of ELLs.

Accommodations Mapped to ELP Levels and Student Profiles

To increase the validity of scores for ELLs on state content assessments given in English, accommodations need to match the linguistic needs of students (Kopriva et al., 2007). ELLs' needs differ across levels of English language proficiency and specifically across the domains of reading and writing in the language(s) spoken by the student. To address this issue, the expert panel mapped accommodations to ELP levels. These mappings contained in the Best Practices study can guide SEAs to select accommodations for ELLs at different levels of English language proficiency and with different levels of literacy in English and the native language.

Accommodations Removed During the Screening Process

The expert panel identified 27 accommodations reviewed in the Descriptive Study of SEA policies that failed to meet the ELL-responsiveness criterion. Because accommodations were originally implemented to support students with disabilities, many states have not distinguished between accommodations for ELLs and students with disabilities. Some items currently listed by

many states as accommodations were considered *test administration practices* by the expert panel, in concurrence with current best practice in five states. That is, these practices are unlikely to reduce construct-irrelevant variance due to language but may be useful logistical adjustments necessary to support the administration of accommodated tests. Some of the accommodations reviewed raised concerns among the expert panel regarding potential threats to validity. These included unscripted accommodations and commercial dictionaries. Scripting accommodations can prevent variations in the presentation of test items that might provide either undue assistance or hinder a student's access to the intended meaning. If commercial dictionaries are allowed, they should not include definitions, explanations, pictures or examples that might afford an unfair advantage to students. If states wish to use dictionaries, these should be vetted to ensure they do not include the answers for particular test items.

Recommendations

The expert panel concluded with recommendations for states to consider when refining policy for the accommodation of ELLs. Recommendations from the Best Practices Study build on the emerging research base and, together with recommendations from the Descriptive Study, provide the basis for the development of the *Guide for Refining State Assessment Policies for Accommodating English Language Learners*.

- Screen accommodations for ELL-responsiveness.
- Standardize and clearly describe accommodations.
- Distinguish between test administration practices and accommodations.
- Offer accommodations for ELLs at each English language proficiency level
- Offer accommodations for ELLs with different levels of literacy in English and the native language.

Introduction

As the standards and accountability movement has gained momentum over the past two decades, the effort to include all students, including English language learners, has also solidified. The Elementary and Secondary Education Act (ESEA) reauthorizations through the 1995 Improving America's Schools Act, and the 2001 No Child Left Behind Act (NCLB) require state, district, and local educational systems to share responsibility for the academic achievement of all students, and specifically require the inclusion of English language learners (ELL) in state assessment systems.

In meeting the inclusion provisions for ELLs in state assessment systems, it is important to ensure the meaningful representation of what students know and can do. The provision of testing accommodations has been a primary strategy states have used to enable ELLs to participate in state assessments as well as to increase the validity and reliability of test scores. Accommodations for ELLs involve changes to testing procedures, testing materials, or the testing situation in order to allow ELLs to participate meaningfully in assessments. Effective accommodations for ELLs address the unique linguistic and socio-cultural needs of the students without altering the test construct.

Accommodated scores should be sufficiently equivalent in scale that they can be pooled with unaccommodated scores.

Accommodations for ELLs involve changes to testing procedures, testing materials, or the testing situation in order to allow ELLs to participate meaningfully in assessments. Effective accommodations for ELLs address the unique linguistic and socio-cultural needs of the students without altering the test construct. Accommodated scores should be sufficiently equivalent in scale that they can be pooled with unaccommodated scores.

State Education Agency (SEA) staff represented at the LEP Partnership meeting in October 2006 agreed that it would be beneficial to collaborate in the refinement of state assessment policies that include test accommodations allowed for ELLs. In response to this recommendation, ED charged The George Washington University Center for Equity and Excellence in Education (GW-CEEE) to develop a Guide for SEAs to use to improve state assessment policies for accommodating ELLs. As a foundation for the Guide, the GW-CEEE research team designed two studies, the Descriptive Study and the Best Practices Study. For the Descriptive Study the research team reviewed state assessment policies and examined the number and types of accommodations specified for ELLs (Shafer Willner, Rivera, & Acosta, 2008). This study of best practices applied a Delphi technique to obtain consensus from an expert panel about which accommodations identified in the descriptive study were Ell-responsive. Because of the scant research base on accommodations for ELLs and the limited amount of work related to mapping accommodations to English language proficiency level (ELP), the expert panel vetted and mapped the ELL-responsive accommodations to ELP levels and to selected student background variables. Members of the panel were selected from experts knowledgeable about research, policy and practice in the areas of assessment, psychometrics, language testing, second language acquisition and instruction of ELLs.

Review of Literature

To respond to the accountability requirements of ESEA, every state has formulated policies for including and accommodating ELLs. These policies have developed without the benefit of an extensive research base. Much of the work on accommodations in large-scale assessments has traditionally focused on students with disabilities, and many state assessment policies have historically drawn from this framework when specifying accommodations for ELLs (Rivera, Collum, Shafer Willner & Sia, 2006). The challenge of making the content of the test accessible to students is different for ELLs than for students with disabilities. Rivera et al. (2006) recommended that state policies offer accommodations that are ELL-responsive – i.e., provide assistance in overcoming the linguistic barriers that prevent ELLs from demonstrating the academic knowledge and skills tested. Without adequate accommodations, ELL test scores cannot accurately reflect what students know and can do (LaCelle-Peterson & Rivera, 1994).

Construct-irrelevant variance is one type of systematic error introduced when a student takes an academic test in a language in which she or he is not yet proficient. Messick (1989) defines construct-irrelevant variance as measurement error due to extraneous information unrelated to the construct being assessed. Construct-irrelevant variance for an ELL can be introduced in the form of language, graphics or cultural references that are not part of the construct assessed. Because ELLs are still in the process of attaining proficiency in English and may not have the cultural knowledge required, complex language and cultural references used in a test item can create a barrier to understanding and responding. For example, in a mathematics test, if an item intended to measure the student's understanding of fractions contains unfamiliar grammatical structures, vocabulary or syntax unrelated to the construct measured, the linguistic complexity in the test item could prevent an ELL from fully understanding the item and demonstrating what he or she knows about fractions.

Second language acquisition research contributes a theoretical framework for analyzing the linguistic challenges faced by ELLs taking large scale assessments. Compared with native English speaking peers, for example, ELLs have not yet automatized their English language processing skills, placing them at a disadvantage when taking a test in English. Particularly at the early stages of language acquisition, research indicates that ELLs encode and decode text in English at slower processing speeds than in their native language(s). In addition, both short-term and working memory may be significantly taxed during second language processing. (See Rivera et al. 2006 for a review.)

To reduce construct-irrelevant variance and to address the needs of ELLs taking large-scale assessments, an effective accommodation must provide linguistic support to help the student access the content of the test (Rivera, et al., 2006). Accommodations must also meet standards of validity and reliability established in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). That is, an accommodation cannot alter the construct being assessed or provide undue assistance in answering the test item (Elliott, Kratochwill, & Schulte, 1998; Koenig & Bachman, 2004). For example, on a math item asking the student to calculate the average time for three runners, it would be appropriate to provide a

glossary with a definition of the term “runner” but not of the term “average.” Defining the term “runner” would help an ELL access the meaning of the test item without giving away the answer, but an explicit definition of the term “average” might provide an unfair advantage. In sum, the challenge of appropriately accommodating ELLs is to ensure that test takers have a fair opportunity to demonstrate what they know and can do without “compromising the validity and other technical aspects of the test” (Rivera et al., 2006, p. 7).

ELL-Responsive Taxonomy

Rivera et al. (2006) developed a taxonomy to distinguish ELL-responsive accommodations from those intended for students with disabilities. This taxonomy classifies accommodations according to whether they provide direct or indirect linguistic support. Both forms of accommodation address ELLs’ linguistic needs, helping these students access the academic construct being measured by the assessment. Accommodations providing direct linguistic support involve adjustments to the language of the test. Such accommodations can be provided in the student's native language or in English. Indirect linguistic support accommodations are defined as adjustments to the conditions under which ELLs take the test. In the recent study of state assessment policies, the ELL-responsive taxonomy was updated (Shafer Willner et. al, 2008). A description of accommodations in each category in the updated taxonomy is found in Table 1.

Effectiveness of Accommodations for ELLs

The body of literature examining the effectiveness of accommodations for ELLs has begun to emerge only within the past decade, concurrent with growing interest in the appropriate inclusion of this population in large scale assessments. Research on accommodations for ELLs to date is limited by the number of accommodations studied, the generalizability of findings, and the meager number of studies per accommodation type. Moreover, studies of accommodations for ELLs have generally not controlled for important student background variables including level of English language proficiency (ELP), and literacy in the native language, and language of instruction.

To date, the number of studies of accommodations provided to ELLs is sparse in comparison to similar studies for students with disabilities. For example, Francis, Rivera, Lesaux, Kieffer (2006) identified 11 empirical studies conducted between 1999 and 2005, and Pennock-Roman and Rivera (2007) identified two additional experimental studies for a total of 13 studies conducted between 1990 and 2007. This is in contrast to the number of studies of accommodations conducted to address students with disabilities. In a one-year period from 2005 to 2006, researchers conducted 32 experimental studies (Zenisky & Sireci, 2007, p. 4).

Of the 13 empirical studies of ELLs, eight accommodations have been examined including English and dual language dictionaries and glossaries, plain English tests, side-by-side dual language tests, translated (Spanish) tests, and extended time. In their meta-analysis, Francis et al. (2006) concluded that many of the accommodations had little to no consistent effect, and none was sufficient to “level the playing field” for ELLs. Only English language dictionaries had a consistent and significant overall positive effect across studies. The use of translated (Spanish)

Table 1

Direct and Indirect Linguistic Support Accommodations for English and Native Language

Direct Linguistic Support Accommodations	
English	Native Language
<i>Plain English</i> consists of test items and/or test directions for which linguistic complexity has been reduced while maintaining the level of difficulty of the test construct. Plain English text is characterized by linguistic structure(s) and vocabulary that avoid ambiguity, colloquialisms, or multiple meanings. Plain English is also referred to in the research literature and/or in state assessment policies as “modified English,” “simplified English,” “simplification,” or “plain language” (Abedi & Sato, 2007; Miles, Rivera, & Stansfield, 2000; Rivera & Stansfield, 2004).	<i>Written translation</i> is the rendering of all or part of an English-language assessment into a second language. Accommodations in this category include written translation of test directions, side-by-side dual language versions of the test, or translated versions of entire tests.
<i>Scripted oral English</i> includes reading aloud and repeating test items or directions from a script and/or presenting the text to the test-taker through an audio recording or CD. The learner simultaneously has access to the written text.	<i>Scripted oral translation</i> involves reading aloud a professionally translated script of translated test items and/or directions or presenting the translation through an audio recording or CD.
<i>Oral Response in English</i> includes accommodations that allow students to answer test items orally in English. Oral responses are tape-recorded and/or scribed and entered onto the student’s test form by the test administrator.	<i>Response in native language</i> involves allowing students to respond either orally or in writing in their stronger language. Native language responses may either be scored in the native language or translated into English prior to scoring.
<i>English reference materials</i> include English dictionaries and glossaries provided in print or electronically. A dictionary defines words. The types of English dictionaries used to accommodate ELLs include standard dictionaries, learners’ dictionaries, and customized dictionaries. A learner’s dictionary is designed specifically for ELLs and defines words in plain English. Like some standard English dictionaries, learners’ dictionaries also give examples of usage and may provide synonyms. A customized dictionary refers to a dictionary that has been altered or specially compiled for a given context. It may refer to a learner’s dictionary in which language has been simplified specifically for ELLs. A customized dictionary also may contain a specialized list of standard dictionary definitions compiled for a particular assessment and containing words relevant to that assessment. English glossaries are specialized lists of words with definitions or explanations customized to fit the perceived needs of the test taker. Glossaries may use simplified English.	<i>Dual language reference materials</i> consist of dictionaries and glossaries provided in print or electronically in both English and a second language.
<i>Clarification in English</i> involves the provision of oral explanations of text considered potentially difficult for ELLs to access. The reformulated input, which is sometimes offered in sheltered English, is expected to be more easily understood and manageable. Clarification differs from scripted oral English in that instead of reading from a script, the test administrator provides the explanation or clarification on-the-fly. Examples of clarification accommodations include “clarify/explain test directions in English,” and “simplify test directions.”	<i>Sight translation</i> is the oral, on-the-fly rendering of test directions, items, or both from English into a student’s native language. This type of accommodation differs from scripted oral translation in that instead of reading from a script, the test administrator (who is typically competent in the language of the translation) orally translates as he or she reads. This “on-the-fly” interpretation also distinguishes sight translation from written translation (Stansfield, 2008).
Indirect Linguistic Support Accommodations	
<i>An indirect linguistic support accommodation</i> consists of extending the time limits of a test to facilitate ELLs’ language processing. Extended time may be provided as a stand-alone accommodation or in combination with one or more direct linguistic support accommodations. The use of dictionaries, glossaries, side-by-side dual language tests, oral accommodations (e.g., read-alouds or oral translations), and response accommodations all require extended time for handling the additional materials or adjustments.	

tests and dual language word-to-word dictionaries also were found to be effective in some studies, although outcomes varied. The researchers cautioned that dictionaries are appropriate only if students are familiar with how to use them and concluded that, to benefit from translated assessments, students need to have received recent native language instruction in the content tested.

By refining the unit of analysis to account for student English language proficiency level and controlling for time restrictions on tests, Pennock-Roman and Rivera's (2007) meta-analysis identified six accommodations with positive effects for ELLs at different levels of English language proficiency. Effective direct linguistic support accommodations included pop-up English language dictionaries/glossaries, side-by-side dual language (Spanish-English) tests, and English dictionaries/glossaries. At lower English language proficiency levels and for students who received Spanish instruction in the content assessed, translated (Spanish) assessments were found to be effective. For students at intermediate levels of English language proficiency instructed in English, a plain English version of the test was found to be effective. Of the two indirect linguistic support accommodations studied - extended time and small group administration - only extended time had a significant effect. Extended time was found to be somewhat effective alone, but more effective in combination with a direct linguistic support accommodation (e.g., a dictionary or glossary).

Although these studies offer some insight into potentially effective accommodations for ELLs, the researchers advise that the research is insufficient and should be interpreted cautiously. Another limitation of the research is the fact that the majority of the experimental studies of accommodations used repurposed items from the National Assessment of Educational Progress (NAEP), not operational or field test items from state assessments. These studies (1) varied in how the same accommodation was constructed and administered; (2) varied in the ages and grade levels administered; and (3) varied in the extent to which student background variables such as level of English language proficiency were controlled.

Considering the limited research base, there is significant need for more research on accommodations for ELLs to be able to make generalizations about the effectiveness of specific accommodations in state assessments. More types of ELL-responsive accommodations need to be investigated and replication studies conducted.

Mapping Accommodations to a Heterogeneous ELL Population

Other researchers also acknowledge the need to attend to the diversity of the ELL population when considering accommodations. Students who are learning English as a second language vary widely on a range of factors such as level of English language proficiency, level of literacy in English and the home language, age, continuous years of formal schooling, language(s) of instruction, and type(s) of language support program (Abedi, 2004; Martiniello, 2007; Solano-Flores, 2006; Solano-Flores & Trumbull, 2008).

Solano-Flores and colleagues (2006) and Solano-Flores and Li, (2006) found that measurement error in tests can be attributed to the combined effects of the test item, individual student differences, and language factors. Findings from research (Solano-Flores, 2006) investigating these interactions indicated that the academic and language development of ELLs tends to be

uneven and that even students at the same English language proficiency level may vary substantially in terms of the amount of measurement error in tests, regardless of whether the assessment is given in English or the native language. This variation was attributed not only to differences in content knowledge but also to varying strengths and weaknesses in students' language proficiency as well as the differing level of linguistic challenge experienced by each student for particular test items.

The mapping of accommodations to students' specific needs also has recently received attention in the research literature. Kopriva, Emick, Hipolito-Delgado, Porfirio, and Cameron (2007), for example, observed that the test scores of ELLs who received inappropriate or incomplete accommodations showed no significant effect, while ELLs who received accommodations that were matched to their particular challenges performed at a significantly higher level. Language of instruction is also a factor in whether ELLs can benefit from accommodations in either English or their native language (Hofstetter, 2003). Francis and his colleagues' (2006) meta-analysis of research on accommodations for ELLs found that students' level of oral and written English (and native) language proficiency impacted the effect size of some accommodations, in particular, customized English dictionaries or glossaries, bilingual dictionaries or glossaries, and native language test booklets. The authors concluded that "the choice of bilingual or native language assessments as an accommodation for ELLs must take into account the students' oral proficiency and literacy in their native language, as well as the language in which they have been instructed" (p. 28). However, the heterogeneity of the ELL population presents a complex puzzle for policy makers as well as school-based decision makers who face the task of selecting appropriate accommodations for individual students.

Reviews of Assessment Policies for ELLs

Given that research on the impact of including ELLs in state assessments is emergent, states have struggled to keep up with the requirements to develop assessment policies that address the needs of ELLs. To examine the responsiveness of state policies to the needs of ELLs taking state assessments, GW-CEEE conducted a series of state assessment policy reviews over the last decade (Rivera et al., 1997; Rivera, Vincent, Hafner, & LaCelle-Peterson, 1997; Rivera et al., 2006; Rivera, Stansfield, Scialdone, & Sharkey, 2000; Shafer Willner et al., 2008). The researchers concluded that states offer many accommodations for ELLs that have not yet been studied. Moreover, it was found that a substantial number of the accommodations offered for ELLs lacked a clear justification for use with this population of students. Rivera et al. (2006) for example, found that ELLs and students with disabilities (SDs) tended to be grouped together within state policies as "special needs students," "at-risk students," or "special populations." Many state policies used combined lists of accommodations which did not distinguish between those accommodations appropriate for ELLs and those appropriate for other students. Of the 75 accommodations identified in 2000-01 state policies, only 44 were related to the linguistic needs of ELLs. The remaining 31 accommodations were relevant only to students with disabilities and included such accommodations as the use of Braille and special lighting conditions. A majority of states' policies (27) were organized according to a taxonomy commonly used for students with disabilities that included four categories; (1) timing/scheduling; (2) setting; (3) presentation; and (4) response (Thurlow, Ysseldyke, & Silverstein, 1993). These categories however fail to draw attention to the language issues relevant to ELLs.

State assessment policies affect the types of accommodations individual ELLs receive at the local level. Shafer Willner, Rivera, and Acosta (2007) examined how school-based decision makers selected accommodations for ELLs in NAEP. Many school-based decision makers reported that they based their inclusion and accommodation decisions largely on guidance provided by state assessment policies. Findings indicated that many decision makers, following their understanding of state policy guidance, either did not provide accommodations for ELLs or selected accommodations for these students using criteria designed for students with disabilities.

Compared with a previous review of 2000-01 policies, the Descriptive Study conducted in conjunction with the current study indicated that seven years later, assessment policies continued to vary widely in terms of the extent to which they were responsive to the needs of ELLs. The review identified a combined total of 104 accommodations allowed for ELLs across the 51 state assessment policies. Only 40 of these met the criteria for ELL-responsiveness. The overall number of accommodations offered that appeared unrelated to ELL needs is an indicator that understanding of these issues continues to vary widely across states. States used many different approaches with varying rationales for accommodating ELLs, in many cases resulting in either an overly broad or overly narrow set of accommodations (Shafer Willner et al., 2008). These findings suggest that states have made some progress, but much work remains for all SEAs to develop state assessment policies that include one hundred percent ELL-responsive accommodations and clear guidance on how to implement the policy.

Method

The purpose of the study is to extend understanding of ELL-responsive accommodations. Because the research base on the validity and effectiveness of accommodations for ELLs remains scant, the research team used a Delphi process, a methodology that relies on professional judgment.

The Delphi was popularized by the RAND Corporation in the 1960's as a means of eliciting judgment from a panel of experts when only partial information is available (Dalkey & Helmer, 1963; Helmer, 1966). The method involves an iterated series of formal questionnaires in which individual responses are collected anonymously, then compiled and reported back to panel members in summary form. Panel members review the feedback and vote again on the same questions until the group achieves a degree of consensus. Dalkey (1969) suggests that the anonymity of responses, iteration and controlled feedback, and statistical compilation of group responses are important characteristics that help reduce the bias often introduced in less systematized, face-to-face interactions. Empirical research has indicated that the Delphi is a reliable means of improving the accuracy of group estimates (Dalkey, 1969).

For this study an expert panel consisting of nationally recognized researchers and state and local practitioners and policy makers was formed. Members of the panel were selected to represent a variety of areas of expertise as well as a range of perspectives (e.g., language testing, linguistics, second language acquisition, policy and practice in the areas of assessment, accommodations and instruction of ELLs). The panel participated in a series of online surveys and face-to-face discussions to (a) define accommodations for ELLs; (b) screen an initial list of accommodations for ELLs currently offered by states and to identify those which did not meet criteria for ELL-

responsiveness; (c) select and rank-order ELL-responsive accommodations; and (d) map accommodations to ELL background variables including English language proficiency and level of literacy.

The Delphi Study was conducted in tandem with and builds on the work of the Descriptive Study (Shafer Willner, Rivera, & Acosta, 2008). As part of the Descriptive Study state policies were examined and 104 accommodations for ELLs identified. Initial screening by the GW-CEEE research team resulted in a list of 66 potentially ELL-responsive accommodations for review by the expert panel. The majority of these accommodations are intended for use with state content assessments given in English.

Initial screening conducted prior to initiating the Delphi process by the research team consisted of removing accommodations clearly designed for students with disabilities and not for ELLs. Further refinements made to the list of accommodations included (a) combining like accommodations; (b) increasing the clarity of the language used to describe individual accommodations; and (c) standardizing the language of the accommodations to the extent possible within the research team's interpretation of the states' intent. Accommodations were organized according to Rivera et al.'s (2006) ELL-responsive taxonomy, which includes direct linguistic support accommodations in English or the native language, and indirect linguistic support accommodations.

The Delphi process was initiated through an introductory web conference to familiarize the expert panel with the process and provide background about the project. This was followed by three rounds of online questionnaires, with each questionnaire re-administered three times. The results of each previous iteration, including participants' rationales for their decisions, were shared with the expert panel for consideration in the following iteration. The process ended with a two-day face-to-face meeting in Arlington, VA. Table 2 outlines the procedures implemented for this study.

During Round 1 of the questionnaires, the expert panel was provided with the list of 66 accommodations. Instructions to the expert panel were to identify accommodations that (a) do not address the linguistic needs of ELLs, or (b) are a threat to the validity of reading/language arts and mathematics assessments.

In Round 2 of the questionnaires, the expert panel considered the potential effects of each of the remaining accommodations. The accommodations were rank ordered according to those most likely to benefit ELLs taking state mathematics assessments. This content area was selected because it was expected to produce the most inclusive list of ELL-responsive accommodations that would not threaten the test construct.

Prior to administering the questionnaires for Round 3, the GW-CEEE research team developed a customized three-level English language proficiency (ELP) rubric (shown in Table 4) to describe the academic language skills students at beginning, intermediate and advanced ELP levels can

Table 2
Summary of Delphi Procedures

Activity	Procedure
Web Conference	Introduce the project and the methodology
Round 1 of online questionnaires	Screen accommodations to identify those that a) are not ELL-responsive, or b) threaten validity
Round 2 of online questionnaires	Rank order accommodations according to the extent to which they are likely to increase the performance of ELLs on state content tests.
Round 3 of online questionnaires	Map each accommodation to 3 levels of English language proficiency (ELP) + selected student background variables in the native language.
Face-to-Face Meeting	a) Agree on list of ELL-responsive accommodations; b) analyze appropriateness of each accommodation for specific academic content assessments; c) map accommodations to ELP levels; d) generate a definition of an accommodation for ELLs

demonstrate in the context of academic content assessments. This specialized ELP rubric is based on indicators found in many current ELP standards such as the Teachers of English to Speakers of Other Languages (TESOL, 2006) and the World-Class Instructional Design and Assessment Consortium (WIDA, 2007). The ELP levels correlate with a new generation of ELP tests that are supported by federal legislation (NCLB) and are designed to measure the degree to which ELLs have acquired academic English (Abedi, 2007). For the third and final round of questionnaires, the expert panel was instructed to refer to the ELP rubric to map the list of accommodations from the previous rounds to the three ELP levels.

Data from the online surveys were compiled and reviewed by the expert panel at a final two-day meeting. At the meeting the expert panel (a) agreed upon the list of ELL-responsive accommodations; b) analyzed the appropriateness of each accommodation for specific academic content assessments; c) refined the mapping of accommodations to ELP levels and student background variables; and d) generated a definition of an accommodation for ELLs.

Limitations of the Study

The results of this study are based on professional opinion of accommodations for ELLs found in state assessment policies and not on empirical analyses. With the exception of the eight accommodations with a research base, the expert panel relied predominantly on its expertise in language testing, experience in administering content assessments to ELLs, and expert knowledge of state assessment policies to select, rank order, and map ELL-responsive accommodations.

Linstone, a veteran practitioner of the Delphi Method, alerts researchers to several common pitfalls (Linstone, 2002). Among those of relevance to the current study is the natural human tendency to reduce uncertainty, particularly when dealing with complex systems or information that may be perceived as counter-intuitive. Linstone further cautions that even groups with

unquestionably high levels of expertise are not invulnerable to human error. Thus it is quite possible for a Delphi process to systematically reproduce error. In general, members of expert panels may be subject to biases that tilt them toward either overly optimistic or overly pessimistic predictions, which in the current study could have affected judgments in applying the operational definition of an accommodation either too liberally or too strictly. In sum, policy changes based on these results should be made with ordinary caution and be subject to periodic review as new research emerges.

Results of the Delphi Process

The results section is divided into three parts. The first part presents the results of the Delphi process in identifying ELL responsive accommodations as well as the rationales for arriving at this final list. The second part discusses the mapping of these accommodations to different ELP levels. A final part includes a discussion of the rationales for eliminating accommodations, including accommodations that were not considered ELL responsive and accommodations that might threaten the validity of mathematics and reading/language arts tests.

ELL-Responsive Accommodations

As an outcome of the Delphi process, the expert panel contributed to the refinement of the operational definition of an ELL-responsive accommodation. Rivera et al.'s (2006) definition of an ELL-responsive accommodation focused on “assistance in overcoming linguistic barriers to enable students to access the content of the assessment and demonstrate what they know and can do” (p. 48). For this study, the definition was operationalized as an accommodation that is likely to *reduce construct-irrelevant variance due to language*.

In applying this operational definition to the accommodations identified in state policies, the expert panel was instructed to first consider what is known from the existing empirical research. For accommodations not yet investigated, the panel was instructed to use professional judgment based on second language acquisition theory, language testing, and knowledge of practice.

The expert panel concluded that 29 accommodations met the operational definition of ELL-responsive. Table 3 contains the categorized list of accommodations, with English accommodations on the left and native language accommodations on the right. Categories of direct linguistic support accommodations include plain English, written translation, reference materials (English and dual language), scripted oral English, scripted oral translation, clarification in English, sight translation, and response in English and the native language. Only one indirect linguistic support accommodation – extended time – remained in the list after screening. Accommodations supported by at least one empirical study are indicated with an asterisk.

Accommodations Mapped to ELP Levels and Student Profiles

An additional task of the expert panel was to map accommodations to English language proficiency (ELP) levels. For purposes of this study, the mapping was conducted for accommodations in both English and the native language to tests given in English. The objective

Table 3
ELL-Responsive Accommodations

		Direct Linguistic Support Accommodations	
		ENGLISH	NATIVE LANGUAGE
WRITTEN	<i>Plain English</i> Provide plain English test *		<i>Written translation</i> Provide translated test * Provide side-by-side dual language test * Provide test directions in native language
	<i>English reference materials</i> Provide customized English glossary		<i>Dual language reference materials</i> Provide customized dual language glossary * Provide customized dual language pop-up electronic glossary * Provide commercial word-to-word dual language dictionary *
			<i>Written response</i> Allow student to respond in writing in native language
ORAL	<i>Scripted oral English</i> Play audio tape/CD of test items Read aloud test items Read aloud test items from plain English script Repeat test items Read aloud test directions Play audio tape/CD of test directions Repeat test directions		<i>Scripted oral translation</i> Play audio tape/CD of test items in native language Read aloud oral script of test items in native language Play audio tape/CD of test directions in native language Read aloud oral script of test directions in native language
	<i>Clarification</i> Clarify/explain test directions Simplify test directions		<i>Sight translation</i> Translate test directions orally into native language Clarify/explain test directions in native language
	<i>Oral Response</i> Allow student to respond orally in English; scribe response Allow student to respond orally in English; use tape recorder to record test response		<i>Oral response</i> Allow student to respond orally in native language; scribe response in native language Allow student to respond orally in native language; translate response to English
		Indirect Linguistic Support Accommodations	
		Allow extended time*	

* Accommodation supported by at least one empirical study

of this activity was to extend current understanding of the effect of specific accommodations across a range of ELP levels. While research investigating this type of mapping is limited, some researchers have begun to examine this issue in accommodation studies (see, for example, Kopriva, Emick, Hipolito-Delgado, Porfirio and Cameron, 2007; Hoffstetter, 2003).

To analyze the multiple ELP levels, the GW-CEEE team developed and used the ELP rubric shown in Table 4. The rubric describes students at three levels of English language proficiency. Using the rubric, the expert panel was instructed to consider each accommodation and the level of ELP a student would need to benefit from the accommodation. This endeavor was challenging

Table 4
GW-CEEE English Language Proficiency Rubric

Level	When taking an academic test in a <i>targeted content area</i> , ELLs...
Beginning	<ul style="list-style-type: none"> ▪ rely extensively on linguistic, visual, and/or graphic support to access meaning from a limited set of words, phrases, and/or sentences up to paragraph level academic text; ▪ write letters, words, groups of words, sentence fragments, phrases, and/or memorized statements; frequent errors in writing impede communication; ▪ comprehend and write a minimal range of <i>targeted content area</i> vocabulary; ▪ have little to no familiarity with academic English or <i>academic discourse</i>; and ▪ have difficulty understanding basic English vocabulary and sentence structure.
Intermediate	<ul style="list-style-type: none"> ▪ rely on previous knowledge of the topic, and/or linguistic, visual, or graphic support to construct meaning from grade-level academic text up to multiple paragraph level; ▪ write multiple sentences related to a topic; produce writing at paragraph up to multiple paragraph level; errors in writing may hinder communication; ▪ understand and use a growing range of specific and technical vocabulary of the <i>targeted content area</i>; ▪ are familiar with some aspects of the structure of <i>academic discourse</i> necessary to comprehend grade-level academic text ▪ interpret <i>contextualized</i> grade-level academic text literally; and ▪ have difficulty understanding new, unfamiliar, abstract, <i>decontextualized</i> academic concepts, complex sentence structure, or abstract <i>targeted content area</i> vocabulary.
Advanced	<ul style="list-style-type: none"> ▪ can construct meaning on most academic topics from grade-level academic text of varying lengths and degrees of linguistic complexity with the benefit of appropriate linguistic, visual, and/or graphic support; ▪ produce organized, cohesive writing comprised of varying sentence lengths and degrees of linguistic complexity; write multiple paragraphs and complete essays; produce some non-native errors that do not interfere with comprehension; level of writing produced may be below grade-level expectations; ▪ understand and use an expanded command of precise, specialized, technical vocabulary of the <i>targeted content area</i>; ▪ are familiar with most aspects of the structure of <i>academic discourse</i> necessary to comprehend grade-level academic text ▪ understand most abstract, <i>decontextualized</i> grade-level academic text on unfamiliar topics; and ▪ comprehend most grade-level academic text at a sufficient level to function in an academic environment with proficient English speaking peers.
<p>Glossary</p> <p><i>Academic discourse</i> – the language, vocabulary, and linguistic structures used for specific academic purposes.</p> <p><i>Contextualized</i> –situated within a context that activates students’ background knowledge; may include visual and graphic supports.</p> <p><i>Decontextualized</i> –requires readers to rely on written text only.</p> <p><i>Targeted content area</i> – subject area of a state assessment (e.g., reading/language arts, mathematics, and writing).</p>	

because, as shown in Figure 1, not only do students vary across ELP levels, they also vary across important background characteristics such as level of literacy in each language, prior education, and language(s) of instruction in the subject areas to be assessed.

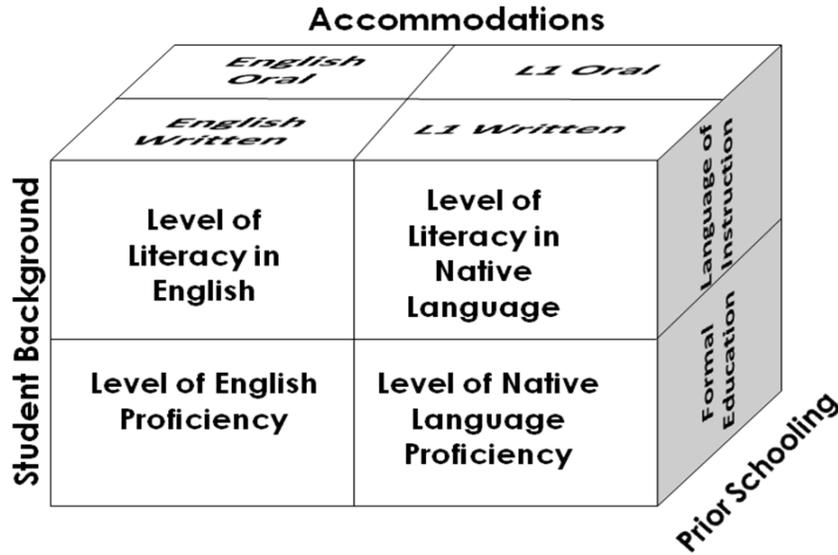


Figure 1. Conceptual Framework for Mapping Accommodations to ELP Levels

In addition, as shown in Table 5, four student profiles were constructed, each representing different combinations of student background variables. Although ideally accommodations in both languages would be mapped to student profiles, the expert panel suggested that most states would not have access to all of the data needed to accomplish this level of analysis. Thus, the student profiles were used for mapping native language accommodations only.

Direct linguistic support accommodations in English were mapped across the three levels of English language proficiency described in the rubric in Table 3. Direct linguistic support accommodations in the native language were mapped to the three ELP levels and to the four student profiles shown in Table 5, creating a four- by three-dimensional crosswalk. The

single surviving indirect linguistic support accommodation, extended time, is independent of language and is therefore included in both mappings. To develop the most extensive list of appropriate accommodations for ELLs, the mapping of accommodations was applied to a hypothetical large-scale state test in mathematics. Table 6 and 7 contain mappings of accommodations in English and in native language. Table 6 maps English language accommodations across the three ELP levels only, and Table 7 maps native language accommodations across ELP levels and to the four student profiles described in Table 5.

Profile	Description
Lit-R	literate in native language and have received recent native language instruction in the content assessed
No-Lit	neither literate in native language nor have received recent native language instruction in the content assessed
Lit-O	older students who are literate in the native language and have received several years of instruction in English for the content assessed
SIFE	students with interrupted formal education (SIFE) who have low literacy skills in both native language and English

Beginning ELP level

ELLs at the lowest levels of English language proficiency tend to experience the greatest need for accommodations on English-language tests but these students are the least able to use them in the judgment of the expert panel. Most accommodations provided in English (Table 6) were considered unlikely to reduce construct-irrelevant variance for ELLs at the lowest levels of English proficiency. Thus, only one direct linguistic support accommodation in English – *read aloud test items from a plain English script* – was highly rated (indicated with a dark circle on the table) for the beginning ELP level. *Provide extended time*, the only remaining indirect linguistic support accommodation, was also highly rated. Five other English language accommodations were rated as potentially helpful for beginners by a smaller number of panel members (indicated with an open circle on the tables). In sum, recommended accommodations in English for the beginning level consisted primarily of oral rather than written accommodations, although some panel members felt that tests written in plain English could potentially help more advanced beginners. However, few English accommodations were expected to be effective for “true beginners,” i.e., non-English proficient students taking a test in English.

Table 6
Mapping of Accommodations in English to ELP levels

Accommodation	ELP Levels		
	Beginning	Intermediate	Advanced
<i>1. Plain English</i>			
1.1. Provide plain English test	○	●	●
<i>2. English reference materials</i>			
2.1. Provide customized English glossary		●	●
<i>3. Scripted oral English</i>			
3.1. Play audio tape/CD of test items		○	
3.2. Play audio tape/CD of test directions		○	
3.3. Read aloud test items	○	○	
3.4. Read aloud test items from plain English script	●	●	
3.5. Read aloud test directions		●	
3.6. Repeat test items	○		
3.7. Repeat test directions	○		
<i>4. Clarification in English</i>			
4.1. Clarify/explain test directions		○	
4.2. Simplify test directions	○	●	
<i>5. Oral Response</i>			
5.1. Allow student to respond orally in English; scribe response		○	
5.2. Use tape recorder to record test responses		○	
<i>6. Allow extended time</i>	●	●	●
● Likely to reduce construct-irrelevant variance for ELLs at this English language proficiency level.			
○ May reduce construct-irrelevant variance for ELLs at this English language proficiency level.			

Native language accommodations shown in Table 7 were among those expected to be most useful to students at beginning levels of ELP. For students who are literate in their native language and who have received instruction in that language (indicated as Lit-R on the table), written translations or dual language versions of the test were favored.

Table 7
Mapping of Native Language Accommodations to ELP levels and Student Profiles

Accommodation	ELP Levels										
	Beginning				Intermediate			Advanced			
Student Profiles	Lit-R	No-Lit	Lit-O	SIFE	Lit-R	No-Lit	Lit-O	SIFE	Lit-R	No-Lit	Lit-O
<i>1. Written translation</i>											
1.1. Provide translated test	●		○		○						
1.2. Provide side-by-side dual language test	●		●		●		●	○			
1.3. Provide written test directions in native language	●		●								
<i>2. Written response</i>											
2.1. Allow student to respond in writing in native language	○		○	●		○		○			
<i>3. Dual language reference materials</i>											
3.1. Provide customized dual language glossary	○		●		●		○	○	●		○
3.2. Provide customized pop-up electronic glossary			●		○		○				○
3.3. Provide commercial word-to-word dual language dictionary					○		○		○		○
<i>4. Scripted oral translation</i>											
4.1. Play audio tape/CD of test items in native language	○	●	○	●	○	○		○			
4.2. Play audio tape/CD of test directions in native language	○	●	○	●		○					
4.3. Read aloud oral script of test items in native language	○	●	○	●	○	○		○			
4.4. Read aloud oral script of test directions in native language	●	●	●	●		○	○				
<i>5. Sight translation</i>											
5.1. Clarify/explain test directions in native language	●	●	●	●		○		○			
5.2. Translate test directions orally into native language	●	○	●	●		○	○	○			
<i>6. Response in native language</i>											
6.1. Allow student to respond orally in native language; scribe response in native language	●	●	●		○						
6.2. Allow student to respond orally in native language; translate response to English	○	●	○	○		○		○			
<i>7. Allow extended time</i>	●	●	●	●	●	●	●	●	●	●	●

○ Likely to reduce construct-irrelevant variance for ELLs at this English language proficiency level.
● May reduce construct-irrelevant variance for ELLs at this English language proficiency level.

Some panel members suggested that Beginning-level ELLs who are literate in a non-alphabetic script such as Chinese or Russian might benefit from oral response accommodations in the native language as a means of demonstrating what they know. With the exception of older students literate in their native language (Lit-O), providing dual language glossaries and dictionaries for an English language test was not considered particularly useful for students with very limited English comprehension because ELLs are unlikely to benefit when the number of unfamiliar words in a passage is extensive (Albus, Bielinski, Thurlow, & Liu, 2001). For beginning students taking a test in English and not literate in their native language (No-Lit), scripted oral translations and response accommodations were recommended, followed by oral response in the native language. Written accommodations in the native language were considered less useful for students in this group.

Older students literate in the native language who have received recent instruction in English in the content to be tested (Lit-O) differ from other groups in that their knowledge of academic

language may be inconsistent across their two languages. Therefore, the panel favored the use of dual language tests and dual language glossaries for this group of students. Oral accommodations were also recommended by a smaller number of panel members.

For students with interrupted formal education (SIFE), knowledge level across the four domains (reading, writing, listening and speaking) of English and the native language is difficult to predict. At the beginning level of English proficiency the SIFE group will likely not have attained grade level literacy in either language. Thus the panel considered oral presentation of test items (e.g., read aloud test items from a plain English script) in English or the native language more useful than a written translated test. The panel also supported the use of translated test directions (clarify test directions in native language, translate test directions into native language) for this group of students.

Intermediate ELP level

Because ELLs at the intermediate level of ELP have usually developed some literacy in English, these students are expected to benefit from a wider variety of both written and oral options. The need for accommodations at this level would be expected to vary considerably depending on the unique background characteristics of the student as well as the literacy demands of the test. Plain English tests, English or dual language glossaries, read-alouds and extended time were considered useful. A dual language test and dual language reference materials were favored for students who are literate in the native language (Lit-R and Lit-O).

Due to the diversity of needs that may occur within the SIFE group, the consensus of the panel was to offer these students more options, with the recommendation that the selection of accommodations, as for any ELL, be thoughtfully considered on the basis of individual student characteristics. The panel observed that some students with low native language literacy may know the content but may not be able to write the response in English. Thus oral accommodations were highly recommended for most students in this group. However, some panel members suggested that students with native language literacy might benefit from written accommodations in the native language.

Advanced ELP level

The need for accommodations was expected to decrease for students at the advanced ELP level. If offered accommodations, the panel observed that these students would be more prepared than students at lower levels of ELP to make use of accommodations such as plain English and English glossaries (Table 6). Dual language glossaries (Table 7) were recommended for students with an advanced ELP level who are literate in the native language and who have received instruction in the content being tested (Lit-R and Lit-O). The panel concluded that students at this level would typically have enough English so they would not need oral translation or response accommodations. The SIFE group was not included in the analysis because it was assumed that students at the advanced ELP level would have moved out of this group.

Prior Experience Requirement

The expert panel questioned the common stipulation that students should always have received a particular accommodation in the classroom prior to allowing it as an accommodation on a test. This prerequisite was considered important for some accommodations, such as dictionaries, but not as important for others such as a plain English assessment, oral English (read-alouds, audiotape/CDs, repetition, or clarification), oral translations, scribed response in English or use of native language tests, or extended time. For example, some students who may not have received recent instruction in their native language might still benefit from scripted oral translation or being allowed to respond in the native language even if the accommodation is not used in instruction. Including the requirement to only use accommodations used in instruction will not work for ELLs because there is no mechanism similar to an IEP to document if an ELL has used the accommodations assigned to him or her as part of instruction.

Accommodations Removed During the Screening Process

Members of the expert panel reviewed the list of accommodations in relation to two types of content assessments; (a) reading/language arts (including both reading and writing constructs), and (b) mathematics. The expert panel was asked to review the list of accommodations using two screening criteria, (a) not ELL-responsive, and/or (b) might threaten the validity of the assessment. The expert panel identified 37 accommodations that met one of the two criteria for elimination. These included 27 accommodations not considered ELL-responsive and another 10 that might threaten validity.

The first subsection that follows describes the expert panel's process for identifying accommodations that met the first criterion (not ELL-responsive). This is followed by an analysis of the second criterion (potential threats to validity). A third subsection describes the panel's analysis of potential threats to validity specifically for reading and writing assessments.

Accommodations not considered ELL-responsive

The expert panel identified 27 accommodations that failed to meet the ELL-responsiveness criterion. As shown in Table 8, some of these accommodations were designed specifically to support students with disabilities and are not relevant to the linguistic needs of ELLs. Accommodations such as closed circuit television or arrows and stop signs, color overlays or place markers used on test forms do not provide linguistic support. Other items listed in previous reviews as indirect linguistic support accommodations failed to meet the operational definition of an ELL-responsive accommodation (Rivera et al., 2006; Rivera et al., 2000). These included items such as facing student during test administration, providing a flexible schedule, etc. Many previously identified indirect linguistic support accommodations however, were considered by the panel to be potentially helpful for facilitating test administration.

Table 8
Accommodations Not Considered ELL-Responsive

1. Administer test in familiar room	14. Ask student to verify understanding of test directions
2. Administer test in location with minimal distraction	15. Cue student to stay on task
3. Administer test in small group	16. Face student during test administration
4. Administer test individually	17. Inform student of remaining time
5. Allow a person familiar to the student to administer the test	18. Provide a flexible schedule
6. Allow student to mark answers in test booklet	19. Provide breaks during test sessions
7. Allow student to move, stand, or pace during assessment in a manner in which others' work cannot be seen and is not distracting to others	20. Provide calculator
8. Allow student to use word processor or typewriter	21. Provide closed circuit television
9. Allow student to point to answers	22. Provide color overlays or templates
10. Allow student to read test aloud to self	23. Provide cues (e.g., arrows and stop signs) on answer form in pencil
11. Ask student to restate test directions in own words	24. Provide multiple sessions for subtests
12. Allow student to use place markers to maintain place	25. Provide preferential seating
13. Allow student to use word processor or typewriter	26. Provide ruler
	27. Provide visual supports

In its review of ELL-responsiveness, the panel discussed the relative merit of accommodating test directions versus test items. Members of the panel indicated that accommodations to test directions do not address strategies for giving students linguistic access to the content of test items. Some accommodations to test directions (e.g., reading test directions aloud) were considered partially helpful and were conserved in the list, while others (“ask student to restate test directions in own words,” or “ask student to verify understanding of test directions”) were eliminated. Thus, while the panel did not remove all accommodations to test directions, it was agreed that accommodations to test directions should not be the target for accommodations. It was agreed that the need for accommodations to test directions could be reduced by making test directions clear and by familiarizing students with the directions and test format prior to the assessment.

Notably, the panel screened out all but one indirect linguistic support accommodation. Administering tests in special settings, with specialized personnel, in small groups or individually are adjustments that might be helpful for increasing students' level of comfort or for facilitating test administration, but these items did not meet the criteria of an ELL-responsive accommodation in the judgment of the panel. Some panel members suggested, for example, that administering the test in a familiar room might reduce ELLs' affective filter. Other items in this category were considered important for facilitating the administration of some kinds of direct linguistic support accommodations (e.g., read-alouds). However, the panel members agreed that these kinds of timing and setting adjustments would not be expected to reduce construct-irrelevant variance due to language.

Extended time was the only indirect linguistic support accommodation that was expected to be effective for ELLs. This accommodation was considered at least partially helpful because processing speed in the second language tends to be slower (Bowles & Stansfield, 2008; Stansfield & Bowles, 2006). Extended time was also considered important for use in combination with direct linguistic support accommodations that require additional time to use or process, such as dictionaries, read-alouds and oral translations. When these accommodations are

provided on tests with restricted time limits, the expert panel expressed concern that the effectiveness of the accommodation may be diminished unless ELLs are provided extra time. Typically, current state academic content tests do not impose restricted time limits. However, schools often attempt to limit tests to a single 45-minute period. To ensure that sufficient time is provided for accommodating students, the panel suggested that state policies may need to continue listing extended time as an accommodation.

Accommodations that might threaten validity

The second criterion for screening out accommodations was operationally defined as the likelihood that an accommodation could alter the test construct, thus threatening the comparability of accommodated test scores. For accommodations that were determined to be ELL-responsive, the panel was instructed to consider issues of validity as the accommodation might be applied to mathematics and reading/language arts assessments. The expert panel relied on their knowledge of psychometrics, existing research on the accommodation (when available), knowledge of second language acquisition, and professional judgment. The expert panel was also instructed to assume that each accommodation was to be implemented within the context of standard testing procedures.

Some accommodations were considered problematic by the expert panel because they could either alter the test construct or provide an unfair advantage. For example, unscripted accommodations in the categories of oral clarification and sight translation can lead to variations in test administration, which can create construct-irrelevant variance in the opinion of the expert panel. Interpreters conducting unscripted oral translations may differ widely in their ability to translate items or instructions. Thus unscripted oral translations of all or part of a test might either provide undue assistance with the answers or hinder performance if the translated version is not equivalent to the English test item. The panel recognized the difficulties most states face in screening, hiring and training qualified translators. Moreover, even when competent translators are available, members of the panel agreed that oral translations should be scripted.

Other kinds of unscripted accommodations were considered troublesome due to their nonstandardized nature and the difficulty of reliable implementation. For instance, administrators who provide examples of an item or task without a script may make a test item easier for some testing sessions, an advantage not shared by students attending other testing sessions. In synthesis, unscripted accommodations carry the risk that a test administrator could either provide undue assistance with the test construct or, conversely, inadvertently hinder students from understanding the item as it was intended. To avoid these kinds of issues, the expert panel recommended that all accommodations be standardized to the extent possible.

Finally, the panel screened from the list of acceptable accommodations commercial dictionaries that contain explanations, definitions, pictures or examples of terminology because they could provide an unfair advantage. Instead the panel recommended that SEAs offering dictionary accommodations develop lists of commercial English or word-to-word dual language dictionaries approved for use with specific assessments.

While all of the accommodations listed in Table 2 were considered appropriate for mathematics assessments, the expert panel judged that some might threaten the validity of reading/language arts assessments. The constructs of reading and writing were considered to have differing validity implications related to accommodations. For *reading*, the panel concurred that the use of read-alouds and audio recordings of reading passages would threaten validity. The panel suggested that it might be appropriate as well to provide a read-aloud or audio recording of multiple choice questions following a reading passage. The panel also recommended that state assessment policies provide explicit guidance regarding when to allow and when to prohibit these accommodations.

Discussion

As states are increasingly held accountable for the performance of all students, the need for effective and valid means of assessing ELLs is ever more urgent. States have relied on accommodations as one of the principle means to increase the validity of ELL test scores. Yet current knowledge about effective accommodations for ELLs is limited. Moreover, because accommodations were originally implemented to support students with disabilities, many states have not distinguished between accommodations for ELLs and students with disabilities. Although some states have attempted to pursue more sophisticated approaches and to include ELL-responsive accommodations in policy, little research exists to guide policy refinements. The results of this study provide critical groundwork to help states refine their policies.

ELL-Responsiveness

Until now, many states have tended to define accommodations for ELLs according to a disabilities framework. The Descriptive Study indicates that states have made uneven progress since a previous review of 2000-01 state assessment policies. Without adequate guidance for accommodating ELLs, states have struggled to define what does and does not “count” as an accommodation for an ELL. Thus some states have offered long lists of accommodations that are not appropriate for ELLs while others have provided very limited choices for accommodating these students (Shafer Willner, Rivera, & Acosta, 2008).

To assist states in identifying accommodations that are likely to address the needs of ELLs taking state content assessments, the expert panel recommended a revised operational definition of an ELL-responsive accommodation. An accommodation can be considered ELL-responsive if it is *likely to reduce construct-irrelevant variance due to language*.

In applying the operational definition to the accommodations extracted from the Descriptive Study the expert panel found that nearly two thirds of the accommodations currently offered to ELLs do not meet the criteria for ELL-responsiveness. The panel observed that a majority of the accommodations offered to ELLs were originally developed for students with disabilities. Many of these accommodations, such as the use of Braille or color overlays would not be expected to reduce construct-irrelevant variance because they do not address the linguistic or cultural needs of ELLs. These accommodations were not considered appropriate to include in assessment policies for ELLs.

The panel also judged that a number of other items traditionally listed as timing/scheduling and setting accommodations were not likely to reduce construct irrelevant variance due to language. Nevertheless, some of these practices were considered important for facilitating test administration. These practices are discussed in a subsequent section.

The expert panel analyzed the relative effectiveness of accommodations to test *directions* vs. test *items*. Based on this analysis, it was concluded that understanding test directions was simply part of proper test administration procedures for all students (American Educational Research Association et al., 1999). Students must be able to understand the directions in order to demonstrate their understanding of a content assessment. Thus, while the expert panel did not go so far as to remove accommodations to test directions they suggested that it is the responsibility of the test administrator to ensure all students understand test directions.

Mapping Accommodations to ELP and Literacy Levels

To increase the validity of scores for ELLs on state content assessments given in English, accommodations need to match the linguistic needs of students (Kopriva et al., 2007). ELLs' needs differ across levels of English language proficiency and specifically across the domains of reading and writing in the language(s) spoken by the student. Considering this diversity, the expert panel concluded that state policies for accommodating ELLs must move beyond a one-size-fits-all approach. The mapping of accommodations to ELP and literacy levels, and as possible, to other relevant background variables helps ensure that test-takers' differing needs are addressed. The literature suggests and the panel agreed that written accommodations are most appropriate for students who are literate in the language of the accommodation, while oral translations may be used for students with fewer literacy skills.

In addition to a consideration of oral versus written modes for accommodations, the expert panel agreed that ELLs differ in their need for support in either English or the native language. It is important to provide some accommodations in students' native languages because these may be among the few accommodations that can be used by ELLs who are recent arrivals at the beginning levels of ELP. In addition, native language accommodations should be allowed for students who have received instruction in their native language in the content area being assessed. Students who have received recent instruction in English, particularly those at the intermediate and advanced ELP levels, are more likely to benefit from accommodations in English.

Avoiding Threats to Validity

Some accommodations raised concerns among the expert panel regarding potential threats to validity. These included unscripted accommodations and commercial dictionaries. Scripting accommodations can prevent variations in the presentation of test items that might provide either undue assistance or hinder a student's access to the intended meaning. If commercial dictionaries are allowed, they should not include definitions, explanations, pictures or examples that might afford an unfair advantage to students. If states wish to use dictionaries, these should be vetted to ensure they do not give away the answers for particular tests.

Some accommodations that are appropriate for mathematics assessments elicited concerns for use with reading assessments. In particular, it might be appropriate to provide a read-aloud or audio recording of multiple choice questions following a reading passage, but not to read aloud the reading passage itself. The panel also recommended that state assessment policies provide explicit guidance regarding when to allow and when to prohibit these accommodations. The expert panel suggested that states collaborate with test developers to carefully screen accommodations so that they do not alter the construct being assessed or provide an undue advantage.

Test Administration Practices

In the Descriptive Study five state policies listed accommodations commonly classified as timing/scheduling and setting (except extended time) as *test administration practices* rather than accommodations. The expert panel for the Delphi Study concurred that these items failed to meet the operational definition of an ELL-responsive accommodation because they are not expected to reduce construct-irrelevant variance. The panel thus recommended removing these practices from the list of ELL-responsive accommodations. For example, administering a test in a location with minimal distraction would not be expected to affect a student's test score, but it would help facilitate reading the test aloud without disturbing other test takers.

Based on this finding, the panel recommended that a distinction be made between accommodations and test administration practices. The panel also concurred that states should consider reclassifying these items as test administration practices, and, as appropriate, integrating them into the directions for administering the test.

Table 9 shows accommodations for which specific test administration practices are recommended. For scripted oral accommodations, sight translations, and response accommodations the panel recommended administering the test in a location with minimal distraction, in a small group or individually. In addition, policy text might include directions for test administrators administering the accommodations to face the student during test administration and to provide preferential seating where the student can hear oral instructions.

In addition to being a stand-alone accommodation, extended time also was recommended as a test administration practice when providing dictionaries and glossaries, scripted oral accommodations, sight translations, and response accommodations.

In conclusion, while the recommended test administration practices were not expected to reduce construct-irrelevant variance, the usefulness of these practices to facilitate test administration was considered a reasonable justification to allow these practices. These may be stated in the policy guiding the administration of accommodations and should also be included in the directions for administering the test.

Table 9
Test Administration Practices for Specific Accommodations

Type of Accommodation	Recommended Test Administration Practice(s)
<i>Oral administration or response</i>	
<ul style="list-style-type: none"> • Scripted oral English • Scripted oral translation • Oral response 	<ul style="list-style-type: none"> • Provide extended time <p>AND</p> <ul style="list-style-type: none"> • Administer test in a separate room OR • Administer test in small group OR • Administer test individually <p>AND (AS NEEDED)</p> <ul style="list-style-type: none"> • Face student during test administration
<i>Reference Materials</i>	
<ul style="list-style-type: none"> • Glossaries and dictionaries • Provide side-by-side dual language test 	<ul style="list-style-type: none"> • Provide extended time

Recommendations

The Delphi process afforded a means of consolidating current knowledge based on the judgment of a panel of experts. Results of this process build on the emerging research base and provide a basis for developing guidance for refinement of state assessment policies for ELLs. The following recommendations are grounded in the findings of the expert panel.

1. Screen accommodations for ELL-responsiveness.
2. Specify accommodations to be used for content assessments (e.g., mathematics, science and reading/language arts).
3. Standardize and clearly describe accommodations.
4. Distinguish between test administration practices and accommodations.
5. Offer accommodations for ELLs at each English language proficiency level
6. Offer accommodations for ELLs with different levels of literacy in English and the native language.

In the sections that follow, each of these recommendations is presented for consideration by states as they engage in refining state assessment policies.

Screen Accommodations for ELL-Responsiveness

Review accommodations allowed in state assessment policies (see Table 2) to ensure that all accommodations offered are appropriate for ELLs. Provided the test construct is not altered, emphasize the accommodation of test items over accommodations to test directions.

Specify Accommodations to Be Used for Content Assessments

State assessment policy should specify the content assessments(s) for which each accommodation is allowed. In addition, the policy should specify any limitations on which parts of the test may be accommodated (e.g., test directions, reading prompt, and/or reading passage).

Carefully select accommodations allowed for reading/language arts assessments based on a consideration of possible threats to validity for the construct being assessed. Generally, plain English, translated or dual language tests, English or dual language glossaries or dictionaries, non-English read-alouds and sight translations of reading passages should not be allowed for stimulus passages used for tests of reading comprehension. However, these accommodations may be appropriate for reading test directions and test prompts.

Standardize and Clearly Describe Accommodations

All accommodations should be standardized to the fullest extent possible so as to avoid the possibility of unduly assisting or hindering performance. Guidance for standardizing specific types of accommodations follows.

- *Plain English* versions of tests should be developed by qualified professionals with experience and/or training in assessment in the content area. The developer should work with a multidisciplinary committee consisting of language specialists and teachers of the content assessed. A number of publications have been published to provide guidance to test developers and states that are interested in developing plain English versions of tests (See, for example, Abedi & Sato, 2007; Miles et al., 2000; Sato, 2007; Rivera & Stansfield, 2004 for additional guidance)
- *Commercial dictionaries* that include explanations and definitions, pictures or examples of terminology should be avoided. States that allow commercial word-to-word dual language dictionaries should consider requiring the use of a vetted list of approved dictionaries that do not provide unwarranted assistance to the student. For example, if a science test item is assessing students' understanding of the concept of osmosis, the dictionary should not include a definition or examples that provide the answer.
- *Scripted oral English*. English read-alouds can be standardized by providing a script or audio tape/CD. Audio recordings should be professionally developed and should be read aloud by a speaker with standard pronunciation and intonation patterns. It is important to specify whether the audio tape/CD is to be played to a small group of students or to individual students and whether the student can control the recording. State policies should also specify that a written version of the test be provided to students to refer to during the oral presentation.
- *Translated tests*. Translations should be developed by professional translators based on standard practice for developing these kinds of tests (See Bowles & Stansfield, 2008 for guidance). The translators should also be specialists or highly experienced in the content of the test. For example, the translator of a science test should have a degree or degrees in

science. The translator should also be an experienced item writer. Back translation is not an efficient way to verify the quality and accuracy of a translation.

- *Scripted oral translation.* The oral translation is best presented to students by a trained administrator competent in the language of the translation or through audio and/or video media. At minimum, the administrator should be highly proficient in both languages, be familiar with the assessment, and have training and/or experience in reading oral scripts of assessments so as to ensure standard administration. Audio recordings should be professionally developed and be read aloud by a “voice over” professional (i.e., a native speaker who is an experienced actor, radio announcer, or other professional who reads aloud with clear standard pronunciation and intonation patterns). The recorded version should be checked to verify that the script has been fully followed and that all words are comprehensible (Stansfield, 2008). The state policy or accommodation implementation guide may specify whether the audio tape/CD is played to a group of students, to individual students, or both. Students should be provided a written version of the translated test to use during the oral presentation.

Distinguish Between Test Administration Practices and Accommodations

Test administration practices should be listed separately from accommodations for ELLs in the assessment policy. Include in the directions for scripted oral accommodations, sight translations, and response accommodations the following test administration practices:

- provide extended time;
- administer test in a separate room, in a small group or individually; and/or
- face student during test administration.

Include a stipulation of extended time in the directions for administering accommodations that require extra reference materials such as glossaries and dictionaries.

Offer Accommodations for ELLs at Each English Language Proficiency Level

Map accommodations for ELLs to ELP levels as defined by the state’s ELP test. Review the policy to ensure accommodations are offered for students at each level. Provide guidance in the policy for assigning accommodations to individual ELLs at each ELP level.

Offer Accommodations for ELLs with Different Levels of Literacy in English and the Native Language

Map accommodations to written and oral language domains in English and the native language. Examine the policy to ensure accommodations are offered to target the needs of ELLs with different literacy levels in English and the native language. Consider providing native language assessments and/or dual language versions of the test in the most frequent languages spoken by the state's ELL students and offering these types of accommodations to students who have received recent instruction in the native language. In the section of the policy providing guidance for assigning accommodations to ELLs, include a consideration of different literacy levels, native language(s), time in the U.S., and language(s) of instruction.

Future Research

As states refine their policies to make them more responsive to the needs of ELLs, new opportunities will arise to expand the research base. This expanded research base will, in turn, inform future refinements to policy. Based on the results of the Delphi process and the Descriptive Study of state assessment policies, we would urge researchers to conduct more experimental studies of accommodations for ELLs, with a focus on both effectiveness and test score comparability. In addition, more research is needed to identify accommodations that would address ELLs' sociocultural needs and to what extent such accommodations would contribute to reducing construct-irrelevant variance due to language.

It is recommended that at a minimum ELP levels and native language literacy be included as covariates in research on specific accommodations. Researchers should use these data, including scores on the state's ELP test, as variables when studying accommodations.

It would be helpful for states to establish monitoring systems to collect data on the accommodations assigned to students as well as the accommodations actually implemented. States should collaborate with researchers by allowing access to data and participating in the design and implementation of relevant studies.

References

- Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, 33(1), 1-14.
- Abedi, J., & Sato, E. (2007). *Linguistic Modification*. Washington, DC: U.S. Department of Education, LEP Partnership.
- Abedi, J. (2007). Part I: Language factors in the assessment of English language learners: The theory and principles underlying the linguistic modification approach. In J. Abedi, & E. Sato (Eds.), *Linguistic Modification* (pp. 1-51). Washington, DC: U.S. Department of Education, LEP Partnership.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bowles, M., & Stansfield, C. W. (2008). *A guide for native language assessments*. Washington, DC: Prepared for the U.S. Department of Education, LEP Partnership.
- Dalkey, N. C. (1969). *The Delphi method: An experimental study of group opinion* No. RM-5888-PR). Santa Monica, CA: Rand.
- Dalkey, N. C., & Helmer, O. (1963). An experimental application of the Delphi Method to the use of experts. *Management Science*, 9, 458-467.
- Elliott, S. N., Kratochwill, T. R., & Schulte, A. G. (1998). The assessment accommodations checklist: Who, what, where, when, why, and how? *Teaching Exceptional Children*, 31(2), 10-14.
- Helmer, O. (1966). *The use of the Delphi Technique in problems of educational innovations* No. P-3499). Santa Monica, CA: Rand.
- Hofstetter, C. H. (2003). Contextual and mathematics accommodation test effects for English language learners. *Applied Measurement in Education*, 1 April 2003 (pp. 159-188.) Lawrence Erlbaum Associates, Inc.
- Koenig, J. A., & Bachman, L. F. (Eds.). (2004). *Keeping score for all: The effects of inclusion and accommodation policies on large-scale educational assessments*. Washington, DC: National Research Council.
- Kopriva, R. J., Emick, J. E., Hipolito-Delgado, C. P., & Cameron, C. A. (2007). Do proper accommodation assignments make a difference? Examining the impact of improved decision making on scores for English language learners. *Educational Measurement: Issues and Practice*, (Fall), 11-20.
- LaCelle-Peterson, M., & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Education Review*, 64(1), 55-75.
- Linstone, H. A. (2002). Eight basic pitfalls: A checklist. In M. Turoff, & H. A. Linstone (Eds.), *The Delphi Method: Techniques and Applications* (pp. 559-571). Newark, NJ: New Jersey Institute of Technology. Retrieved from <http://is.njit.edu/pubs/delphibook/>
- Martiniello, M. (2007). Linguistic complexity and differential item functioning (DIF) for English language learners (ELL) in math word problems. Unpublished PhD dissertation, Harvard University, Cambridge, MA.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement (3rd Ed.)* (pp. 13-103). New York: Macmillan.

- Miles, J. E., Rivera, C., & Stansfield, C. W. (2000). *Leveling the assessment 'playing field': Helping item writers and teachers to develop science test items accessible to English language learners*. Washington, DC: Prepared for the Delaware Department of Education by the Center for Equity and Excellence in Education, Graduate School of Education and Human Development, George Washington University.
- Pennock-Roman, M. & Rivera, C. (2007). *Test validity and mean effects of test accommodations for ELLs and non-ELLs: A meta-analysis*. Washington, DC: National Center for the Improvement of Educational Assessment, WestEd, and the Assessment and Accountability Comprehensive Center.
- Rivera, C., Collum, E., Shafer Willner, L., & Sia Jr., J. K. (2006). An analysis of state assessment policies addressing the accommodation of English language learners. In C. Rivera, & E. Collum (Eds.), *A national review of state assessment policy and practice for English language learners* (pp. 1-173). Mahwah, NJ: Lawrence Erlbaum Associates.
- Rivera, C., Vincent, C., Hafner, A., & LaCelle-Peterson, M. (1997). Statewide assessment programs: Policies and practices for the inclusion of Limited English proficient students. *Practical Assessment, Research & Evaluation*, 5(13) Retrieved from <http://PAREonline.net/getvn.asp?v=5&n=13>
- Rivera, C., Stansfield, C., Scialdone, L., & Sharkey, M. (2000). *An analysis of state policies for the inclusion and accommodation of English language learners in state assessment programs during 1998-1999*. Arlington, VA: The George Washington University, Center for Equity and Excellence in Education.
- Rivera, C., & Stansfield, C. W. (2004). The effect of linguistic simplification of science test items on score comparability. *Educational Assessment*, 9(3), 79-105. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=qeh&AN=BEDI04125753&site=ehost-live>
- Sato, E. (2007). Part II: A Guide to linguistic modification: Increasing English language learner access to academic content. *Linguistic Modification* (pp. 52-87). Washington, DC: U.S. Department of Education, LEP Partnership.
- Shafer Willner, L., Rivera, C., & Acosta, B. (2007). *Decision-making practices of urban districts for including and accommodating English language learners in NAEP: School-based perspectives*. Arlington, VA: The George Washington University Center for Equity and Excellence in Education.
- Shafer Willner, L., Rivera, C., & Acosta, B. (2008). *Descriptive study of state assessment policies for accommodating English language learners*. Prepared for the LEP Partnership, U.S. Department of Education. Arlington, VA: The George Washington University Center for Equity and Excellence in Education.
- Solano-Flores, G. (2006). Language, dialect, and register: Sociolinguistics and the estimation of measurement error in the testing of English language learners. *Teachers College Record*, 108(11), 2354-2379.
- Solano-Flores, G., & Trumbull, E. (2008). In what language should English language learners be tested? In R. Kopriva (Ed.), *Improving testing for English language learners: A comprehensive approach to designing, building, implementing and interpreting better academic assessments* (pp. 169). New York, NY: Routledge.
- Stansfield, C. W. (2008). *A guide to sight translation of assessments*. Washington, DC: LEP Partnership.
- Stansfield, C. W., & Bowles, M. (2006). Test translation and state assessment policies for English language learners. In C. Rivera, & E. Collum (Eds.), *State assessment policy and practice for English language learners: A national perspective* (pp. 176-314). Mahwah, NJ: Lawrence Erlbaum Associates.

Thurlow, M., Ysseldyke, J., & Silverstein, B. (1993). *Testing accommodations for students with disabilities: A review of the literature* (Synthesis Report No. 4). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

