

July 2006

NO CHILD LEFT BEHIND ACT

Assistance from Education Could Help States Better Measure Progress of Students with Limited English Proficiency

GAO reposted the web version of this report on September 20, 2006, to include a link to a Spanish translation of the Highlights page, <http://www.gao.gov/highlights/d061111high.pdf>.

Para ver la versión de “Highlights” en español, <http://www.gao.gov/highlights/d061111high.pdf>, con la fecha del 20 de Septiembre de 2006, oprima aquí.





Highlights of [GAO-06-815](#), a report to congressional requesters

Why GAO Did This Study

The No Child Left Behind Act of 2001 (NCLBA) focused attention on the academic achievement of more than 5 million students with limited English proficiency. Obtaining valid test results for these students is challenging, given their language barriers. This report describes (1) the extent to which these students are meeting annual academic progress goals, (2) what states have done to ensure the validity of their academic assessments, (3) what states are doing to ensure the validity of their English language proficiency assessments, and (4) how the U.S. Department of Education (Education) is supporting states' efforts to meet NCLBA's assessment requirements for these students. To collect this information, we convened a group of experts and studied five states (California, Nebraska, New York, North Carolina, and Texas). We also conducted a state survey and reviewed state and Education documents.

What GAO Recommends

GAO recommends that the Secretary of Education (1) support research on accommodations, (2) identify and provide technical support states need to ensure the validity of academic assessments, (3) publish additional guidance on requirements for assessing English language proficiency, and (4) explore ways to provide additional flexibility for measuring annual progress for these students. Education generally agreed with our recommendations.

To view the Spanish translation of this highlights page, please see <http://www.gao.gov/highlights/d061111high.pdf>.

www.gao.gov/cgi-bin/getrpt?GAO-06-815.

To view the full product, including the scope and methodology, click on the link above. For more information, contact Marnie S. Shaul at (202) 512-7215 or shaulm@gao.gov.

NO CHILD LEFT BEHIND ACT

Assistance from Education Could Help States Better Measure Progress of Students with Limited English Proficiency

What GAO Found

In the 2003-2004 school year, state data showed that the percentage of students with limited English proficiency scoring proficient on a state's language arts and mathematics tests was lower than the state's annual progress goals in nearly two-thirds of the 48 states for which we obtained data. Further, our review of data 49 states submitted to Education showed that in most states, these students generally did not perform as well as other student groups on state mathematics tests. Factors other than student knowledge, such as how a state establishes its annual progress goals, can influence whether states meet their goals.

For their academic assessments, officials in our five study states reported taking steps to follow generally accepted test development procedures and to ensure the validity and reliability of these tests for students with limited English proficiency, such as reviewing test questions for bias. However, our group of experts expressed concerns about whether all states are assessing these students in a valid manner, noting that some states lack the resources and technical expertise to take appropriate steps to ensure the validity of tests for these students. Further, Education's peer reviews of assessments in 38 states found that 25 states did not provide adequate evidence to ensure the validity or reliability of academic test results for these students. To improve the validity of these test results, most states offer accommodations, such as a bilingual dictionary. However, our experts reported that research is lacking on what accommodations are effective in mitigating language barriers. A minority of states used native language or alternate assessments for students with limited English proficiency, but these tests are costly to develop and are not appropriate for all students.

Many states are implementing new English language proficiency assessments in 2006 to meet NCLBA requirements; as a result, complete information on their validity and reliability is not yet available. In 2006, 22 states used tests developed by one of four state consortia. Consortia and state officials reported taking steps to ensure the validity of these tests, such as conducting field tests. A 2005 Education-funded technical review of available documentation for 17 English language proficiency tests found insufficient documentation of the validity of these assessments' results.

Education has offered a variety of technical assistance to help states assess students with limited English proficiency, such as peer reviews of states' academic assessments. However, Education has issued little written guidance to states on developing English language proficiency tests. Officials in one-third of the 33 states we visited or directly contacted told us they wanted more guidance about how to develop tests that meet NCLBA requirements. Education has offered states some flexibility in how they assess students with limited English proficiency, but officials in our study states told us that additional flexibility is needed to ensure that progress measures appropriately track the academic progress of these students.

Contents

Letter		1
	Results in Brief	3
	Background	6
	Students with Limited English Proficiency Performed below Progress Goals in 2004 in Two-Thirds of States, but States We Studied Are Working to Improve Student Academic Performance	13
	Selected States Considered Language Issues When Developing Academic Assessments, but Validity and Reliability Concerns Remain	25
	Most States Implemented New English Language Proficiency Assessments but Faced Challenges Establishing Their Validity and Reliability	35
	Education Has Provided Assistance, but States Reported Need for Additional Guidance and Flexibility	41
	Conclusions	49
	Recommendations for Executive Action	50
	Agency Comments	51
Appendix I	GAO's Group of Experts on Assessing the Academic Knowledge of Students with Limited English Proficiency	53
Appendix II	Determining Adequate Yearly Progress for Student Groups	55
Appendix III	Percentage of Districts Making AYP Goals for Mathematics for Students with Limited English Proficiency	57
Appendix IV	Proficiency Scores on Mathematics Tests for All Students and Students with Limited English Proficiency	58

Appendix V	Enhanced Assessment Consortia Participation	60
Appendix VI	English Language Proficiency Assessments Used in the 2005-2006 School Year, by State	61
Appendix VII	Comments from the Department of Education	63
Appendix VIII	GAO Contacts and Acknowledgments	68
Related GAO Products		69

Tables

Table 1: Selected Provisions from Title I of NCLBA	8
Table 2: Selected Provisions from Title III of NCLBA	10
Table 3: Examples of English Language Proficiency and Language Arts Standards for a Fifth-Grade Student	10
Table 4: Percentage of Elementary Students Scoring at the Proficient Level or Above on State Mathematics Assessment for Selected Student Groups, School Year 2003-2004	19
Table 5: Examples of Issues Relating to Assessing Students with Limited English Proficiency Raised in Education's Peer Review Reports	28
Table 6: Most Frequently Cited Accommodations in 42 States	31

Figures

Figure 1: NCLBA's Requirements for Students with Limited English Proficiency under Title I and Title III	11
Figure 2: School Year 2003-2004 Comparison of Percentage of Students with Limited English Proficiency Who Achieved Proficient Scores in Language Arts and Mathematics with State-Established Progress Goals	15

Figure 3: Percentage of Districts in 18 Selected States Reporting Adequate Yearly Progress Results in School Year 2003-2004 for Students with Limited English Proficiency	17
Figure 4: Use of Native Language and Alternate Assessments for Students with Limited English Proficiency	34
Figure 5: Type of English Language Proficiency Assessment Administered in 2005–2006 School Year	38
Figure 6: Movement of Students In and Out of Limited English Proficient Student Group and Other Student Groups	47
Figure 7: Process for Determining Adequate Yearly Progress for a Student Group	56

Abbreviations

AYP	adequate yearly progress
CELLA	Comprehensive English Language Learning Assessment
ELDA	English Language Development Assessment
ESL	English as a second language
GAO	Government Accountability Office
MWAC	Mountain West Assessment Consortium
NCLBA	No Child Left Behind Act
PA EAG	Pennsylvania Enhanced Assessment Grant
SCASS	State Collaborative on Assessment and Student Standards
WIDA	World-Class Instructional Design and Assessment

This is a work of the U.S. government and is not subject to copyright protection in the United States. It may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



United States Government Accountability Office
Washington, DC 20548

July 26, 2006

The Honorable George Miller
Ranking Minority Member
Committee on Education and the Workforce
House of Representatives

The Honorable Rubén Hinojosa
Ranking Minority Member
Subcommittee on Select Education
Committee on Education and the Workforce
House of Representatives

The Honorable Lynn Woolsey
Ranking Minority Member
Subcommittee on Education Reform
Committee on Education and the Workforce
House of Representatives

The Honorable Raúl Grijalva
House of Representatives

An estimated 5 million children with limited English proficiency were enrolled in U.S. public schools during the 2003-2004 school year, representing about 10 percent of the total school population. They speak over 400 languages, with almost 80 percent of students with limited English proficiency speaking Spanish. These students have difficulties in speaking, reading, writing, or understanding English that interfere with their ability to successfully participate in school. Because of these language barriers, obtaining information on the academic knowledge of these students from an assessment that is valid and reliable (i.e., it measures what it is designed to measure in a consistent manner) presents challenges. As a result, students with limited English proficiency have historically been excluded from statewide assessments, leaving states and districts with little information about how these students are performing academically.

In 1994, the enactment of the Improving America's Schools Act required states to assess these students, to the extent practicable, in the manner most likely to yield accurate information about their academic knowledge. Subsequently, Congress passed the No Child Left Behind Act of 2001

(NCLBA) with the goal of increasing academic achievement and closing achievement gaps among different student groups. Specifically, NCLBA required states to demonstrate that all students have reached the “proficient” level on a state’s language arts and mathematics assessments by 2014. States are obligated to demonstrate “adequate yearly progress” toward this goal each year—that is, they must show that increasing percentages of students are reaching proficient achievement levels over time. Students with limited English proficiency, along with other targeted student groups, must separately meet the same academic progress goals as other students. Further, NCLBA required states to annually assess the English proficiency of these students and to demonstrate that they are making progress in becoming proficient in English. Because these students are defined by a temporary characteristic—unlike other student groups targeted in NCLBA—once a state determines that students with limited English proficiency have attained English proficiency, they are no longer included in the group of students with limited English proficiency, although Education has given states some flexibility in this area.

Given your interest in the academic achievement of these students and the validity and reliability of assessments used to measure their performance, we are providing information on (1) the extent to which students with limited English proficiency are meeting adequate yearly progress goals and what selected states and districts are doing to support the improved academic performance of these students, (2) what states have done to ensure that results from language arts and mathematics assessments are valid and reliable for students with limited English proficiency, (3) how states are assessing English proficiency and what they are doing to address the validity and reliability of these assessment results, and (4) how the Department of Education (Education) is supporting states’ efforts to meet NCLBA’s assessment requirements for these students.

To determine the extent to which students with limited English proficiency are meeting adequate yearly progress goals, we collected school year 2003-2004 state-level data for 48 states, including the District of Columbia. We obtained the majority of our data from state Web sites and, when necessary, contacted state officials for these data. Three states did not publish data in a format that allowed us to determine if students with limited English proficiency had met the state’s adequate yearly progress goals. We also collected additional achievement data for 2003-2004 at the school district level from 18 states. We chose a nonrandom sample of states with the largest percentage of the national population of students with limited English proficiency, states with the largest percentage increases in these students between 1990 and 2000, and

included at least 2 states from each region represented by Education's regional education laboratories (with the exception of one region that included only one state). When district-level achievement data for school year 2003-2004 were not available on a state's Web site or a state had more than 500 districts, we requested the data directly from state officials; 2 states did not respond to our request for these data. We determined that the state and district data were sufficiently reliable for our purposes. We studied 5 states in depth (California, Nebraska, New York, North Carolina, and Texas) to collect detailed information from state and district officials on their assessment practices, efforts to ensure the validity and reliability of their assessments for students with limited English proficiency, and their approaches to improve the performance of these students. These 5 states had relatively large percentages of students with limited English proficiency or had experienced large increases in their populations of these students. In addition, we selected these particular states to ensure variation in geography, types of English language proficiency tests used, and use of different approaches to assessing the content knowledge of this student group. To obtain information on the assessments used by other states, we reviewed accountability workbooks and other documents that states submit to Education, available reports from state monitoring visits conducted by Education, and available peer review reports from 38 states on their assessment and accountability systems. In addition to studying 5 states, we directly contacted officials in 28 states to confirm what English language proficiency assessment they planned to administer in 2005-2006 and to discuss what guidance Education had provided regarding these assessments. We also interviewed officials from major test development companies, from state consortia that are developing English language proficiency assessments, and from Education. To assess state efforts to ensure the validity and reliability of their assessments, we reviewed national assessment standards developed by professional organizations and convened a group of experts to discuss states' efforts to develop and implement valid and reliable academic assessments for students with limited English proficiency (see app. I for more information about these experts). Finally, we obtained information from the 50 states and the District of Columbia on their use of native language assessments using a short e-mail survey. We conducted our review between June 2005 and June 2006 in accordance with generally accepted government auditing standards.

Results in Brief

In school year 2003-2004, state data showed that the percentage of students with limited English proficiency scoring proficient on a state's language arts and mathematics tests was lower than the state's annual

progress goals in nearly two-thirds of the 48 states for which we obtained data. To help these students progress academically, state and district officials in the 5 states we visited reported using a variety of strategies, including training teachers to incorporate language development into academic classes. Further, our review of data 49 states submitted to Education showed that the performance of students with limited English proficiency on states' mathematics assessments for elementary school students was lower than that of the total student population in all of these states but 1. Although the student groups are not mutually exclusive, in most of the 49 states, the performance of students with limited English proficiency was generally lower than that of other groups, such as economically disadvantaged students. Factors other than student academic knowledge, however, can influence whether states and districts meet their academic progress goals for students with limited English proficiency, such as how a state establishes its annual progress goals. To support improved academic progress for these students, district and state officials we spoke with in our 5 study states reported using strategies similar to those considered good practices for all students. In particular, they cited providing teacher training focused on these students, having school leadership focused on their needs and using data to target interventions as key to the success of these students.

For assessments of academic knowledge in language arts and mathematics, we found that our 5 study states have taken some steps to address the specific challenges associated with assessing students with limited English proficiency. Although officials in these states reported taking steps to follow generally accepted test development procedures to ensure the validity and reliability of results from these assessments for the general student population, these assessments may not provide valid results for students with limited English proficiency. Our group of experts expressed concerns about whether all states are assessing these students in a valid and reliable manner, noting that states are not taking all the critical steps needed to do so. Although states have been required to include these students in their assessments since 1994, Education's recent peer reviews of 38 states cited 25 for not providing sufficient evidence on the validity or reliability of results for students with limited English proficiency. In 1 state, for example, procedures to develop test questions did not include an adequate check for language biases. To increase the validity and reliability of assessment results for this population, most states offered accommodations, such as offering extra time to complete the test and using a bilingual dictionary. While most states offered some accommodations, our experts indicated that research is lacking on what specific accommodations are appropriate for students with limited English

proficiency. Our survey of states and review of state documents found that 16 states used statewide native language assessments for some grades and 13 states used statewide alternate assessments (such as portfolios of classroom work) in 2005 to better accommodate certain students with limited English proficiency. While such assessments may improve the validity of test scores, our group of experts noted that developing native language and alternate assessments requires resources and expertise that not all states have. Further, our experts told us that native language assessments may not provide valid results for students who are not receiving instruction in their native language. In addition, developing assessments in all languages spoken by students with limited English proficiency would likely not be practicable for most states.

With respect to English language proficiency assessments, many states are still in the preliminary phases of developing and administering new assessments to measure students' progress in learning English, as required by NCLBA; as a result, complete information on the validity and reliability of these assessments is not yet available. To assess these students in the 2005-2006 school year, 22 states used new assessments developed by one of four state consortia; 8 states used customized, off-the-shelf assessments offered by testing companies; 14 states used off-the-shelf assessments; and 7 states used state-developed assessments. While a few states already had the required English language proficiency assessments in place, others will be administering these assessments for the first time in 2006; as a result, states and test developers are still collecting evidence to document the validity and reliability of the results for most of these tests. An Education-funded study by a national education research organization reviewed the available documentation for the English language proficiency assessments used by 33 states in the 2005-2006 school year and found insufficient documentation on the validity and reliability of results from these assessments.

Education has offered states support and technical assistance in a variety of ways to help them appropriately assess students with limited English proficiency, such as providing training, conducting peer reviews of states' academic assessments, and providing flexibility in assessing these students. However, Education has issued little written guidance to states on developing English language proficiency assessments that meet NCLBA's requirements and on tracking the progress of students in acquiring English. Officials in about one-third of the 33 states we visited or contacted told us that they were uncertain about Education's requirements for the new English language proficiency assessments and that they wanted more guidance. In addition, our group of experts reported that

some states need more assistance to develop language arts and mathematics assessments that provide valid measures of the academic knowledge of this group of students. To support states' efforts to incorporate these students into their accountability systems for academic performance, Education has offered states some flexibilities in how they track progress goals for these students. For example, students who have been in the United States for less than 1 year do not have to be assessed for language arts. However, several state and district officials in the 5 states we studied told us that additional flexibility, such as excluding students from testing for a longer period, is needed to ensure that adequate yearly progress measures accurately track the academic progress of these students.

To help states improve their assessment of students with limited English proficiency, we are recommending that the Secretary of Education (1) support additional research on accommodations, (2) identify and provide additional technical support states need to ensure the validity and reliability of academic assessments for these students, (3) publish more detailed guidance on assessing the English language proficiency of these students, and (4) explore ways to provide additional flexibility with respect to measuring annual progress for these students. In its comments, Education generally agreed with our recommendations.

Background

Students with limited English proficiency are a diverse and complex group. They speak many languages and have a tremendous range of educational needs and include refugees with little formal schooling and students who are literate in their native languages. Accurately assessing the academic knowledge of these students in English is challenging. If a student responds incorrectly to a test item, it may not be clear if the student did not know the answer or misunderstood the question because of language barriers.

Several approaches are available to allow students to demonstrate their academic knowledge while they are becoming proficient in English, although each poses challenges. First, a state can offer assessments in a student's native language. However, vocabulary in English is not necessarily equivalent in difficulty to the vocabulary in another language. As a result, a test translated from English may not have the same level of difficulty as the English version. If a state chooses to develop a completely different test in another language instead of translating the English version, the assessment should measure the same standards and reflect the same level of difficulty as the English version of the test to ensure its

validity. Second, states can also offer accommodations, such as providing extra time to take a test, allowing the use of a bilingual dictionary, reading test directions aloud in a student's native language, or administering the test in a less distracting environment. Accommodations alter the way a regular assessment is administered, with the goal of minimizing the language impediments faced by students with limited English proficiency; they are intended to level the playing field without providing an unfair advantage to these students. Finally, states can use alternate assessments that measure the same things as the regular assessment while minimizing the language burden placed on the student. For example, an alternate assessment can be a traditional standardized test that uses simplified English or relies more on pictures and diagrams. It can also be a portfolio of a student's class work that demonstrates academic knowledge. In either case, studies would be needed to demonstrate that the alternate assessment is equivalent to the regular assessment.

NCLBA Requirements

Title I of NCLBA seeks to ensure that all children have a fair and equal opportunity to obtain a high-quality education and become proficient in academic subjects. It requires states to administer tests in language arts and mathematics to all students in certain grades and to use these tests as the primary means of determining the annual performance of states, districts, and schools. These assessments must be aligned with the state's academic standards—that is, they must measure how well a student has demonstrated his or her knowledge of the academic content represented in these standards. States are to show that increasing percentages of students are reaching the proficient level on these state tests over time. NCLBA also requires that students with limited English proficiency receive reasonable accommodations and be assessed, to the extent practicable, in the language and form most likely to yield accurate data on their academic knowledge. Somewhat similar versions of these provisions, such as reporting testing results for different student groups, had been included in legislation enacted in 1994. One new NCLBA requirement was for states to annually assess the English language proficiency of students identified as having limited English proficiency. Table 1 summarizes some key Title I provisions from NCLBA.

Table 1: Selected Provisions from Title I of NCLBA

State academic assessments	Beginning in the 2005–2006 school year, states must implement annual, high-quality state assessments in reading and mathematics in grades 3-8 and at least once in high school. ^a These assessments must be aligned with challenging state academic standards and must be “consistent with relevant, nationally recognized professional and technical standards for such assessments” and used in ways that are valid and reliable. States must provide for the participation of all students, including those with limited English proficiency.
Academic assessment provisions related to students with limited English proficiency	Students with limited English proficiency are to be assessed in a valid and reliable manner. In addition, they must be provided with reasonable accommodations and be assessed, to the extent practicable, “in the language and form most likely to yield accurate data” on their academic knowledge. In addition, for language arts, students who have been in U.S. schools for 3 years or more generally must be assessed in English.
Adequate yearly progress	States must set annual goals that lead to all students achieving proficiency in language arts and mathematics by 2014. To be deemed as having made adequate yearly progress for a given year, each district and school must show that the requisite percentage of each designated student group, as well as the student population as a whole, met the state proficiency goal (that is, the percentage of students who have achieved the proficient level on the state’s assessments). Schools must also show that at least 95 percent of students in each designated student group participated in these assessments. Further, schools must also demonstrate that they have met state targets on other academic indicators—graduation rates in high school or attendance or other measures in elementary and middle schools. Alternatively, a district or school can make adequate yearly progress through the “safe harbor” provision, if the percentage of students in a group considered not proficient decreased by at least 10 percent from the preceding year and the group made progress on one of the state’s other academic indicators.
Actions when adequate yearly progress not achieved	Schools that receive funding under Title I of NCLBA must take specified actions if they do not meet state progress goals. Specifically, schools that do not make adequate yearly progress for 2 consecutive years or more are identified for improvement and must, among other things, offer parents an opportunity to transfer students to another school and provide supplemental services (e.g., tutoring). Those that miss the annual goals for additional years are identified for successive stages of intervention, including corrective action and restructuring.
State English language proficiency assessments	States must annually assess the English language proficiency of all students with limited English proficiency, measuring students’ oral language, reading, and writing skills in English.

Source: Pub. L. No. 107-110.

^aBeginning in school year 2007-2008, states must implement similar assessments in science.

Accurately assessing the academic knowledge of students with limited English proficiency has become more critical because NCLBA designated specific groups of students for particular focus. These four groups are students who (1) are economically disadvantaged, (2) represent major racial and ethnic groups, (3) have disabilities, and (4) are limited in English proficiency. These groups are not mutually exclusive, so that the results for a student who is economically disadvantaged, Hispanic, and has limited English proficiency could be counted in all three groups. States and school districts are required to measure the progress of all students in meeting academic proficiency goals, as well as to measure separately the progress of these designated groups. To be deemed as having made adequate yearly progress, generally each district and school must show

that each of these groups met the state proficiency goal (that is, the percentage of students who have achieved the proficient level on the state's assessments) and that at least 95 percent of students in each designated group participated in these assessments.

Although NCLBA placed many new requirements on states, states have broad discretion in many key areas. States establish their academic content standards and then develop their own tests to measure the academic content students are taught in school. States also set their own standards for what constitutes proficiency on these assessments. In addition, states set their own annual progress goals for the percentage of students achieving proficiency, using guidelines outlined in NCLBA.¹

Title III of NCLBA focuses specifically on students with limited English proficiency, with the purpose of ensuring that these students attain English proficiency and meet the same academic content standards all students are expected to meet. This title established new requirements intended to hold states and districts accountable for student progress in attaining English proficiency. It requires states to establish goals to demonstrate, among other things, annual increases in (1) students making progress in learning English and (2) students attaining English proficiency. Specifically, states must establish English language proficiency standards that are aligned with a state's academic content standards. The purpose of these alignment requirements is to ensure that students are acquiring the academic language they will need to successfully participate in the classroom. Education also requires that a state's English language proficiency assessment be aligned to its English language proficiency standards. While NCLBA requires states to administer academic assessments to students in specific grades, it requires states to administer an annual English language proficiency assessment to all students with limited English proficiency, from kindergarten to grade 12. See table 2 for summary of key Title III provisions.

¹For more information on differences in state approaches to meeting NCLBA requirements, see GAO, *No Child Left Behind Act: Improvements Needed in Education's Process for Tracking States' Implementation of Key Provisions*, [GAO-04-734](#) (Washington, D.C.: Sept. 30, 2004).

Table 2: Selected Provisions from Title III of NCLBA

State English language proficiency standards	States must establish English language proficiency standards that are aligned with the state’s challenging academic content standards.
Tracking student progress in learning English	States must establish objectives for improving students’ English proficiency in four areas: speaking, listening, reading, and writing. ^a States receiving grants under Title III must establish annual goals for increasing and measuring the progress of students with limited English proficiency in (1) learning English, (2) attaining English proficiency, and (3) meeting adequate yearly progress goals in attaining academic proficiency outlined in Title I.
Actions when annual goals for students with limited English proficiency not met	Districts that receive funding under Title III are subject to certain consequences if they do not meet a state’s annual Title III goals. If a district does not meet the goals for 2 consecutive years, it must develop an improvement plan that addresses the factors that prevented the district from meeting the goals. If a district does not meet the goals for 4 consecutive years, it must modify its curriculum and method of instruction or the state must determine whether to continue to fund the district and require the district to replace all personnel related to the district’s inability to meet the goals.

Source: Pub. L. No. 107-110.

^aTitle I refers to oral language skills, which encompass listening and speaking.

Language arts standards define the academic skills a student is expected to master, while English language proficiency standards define progressive levels of competence in the acquisition of English necessary to participate successfully in the classroom. Examples of standards for English language proficiency and language arts are provided in table 3.

Table 3: Examples of English Language Proficiency and Language Arts Standards for a Fifth-Grade Student

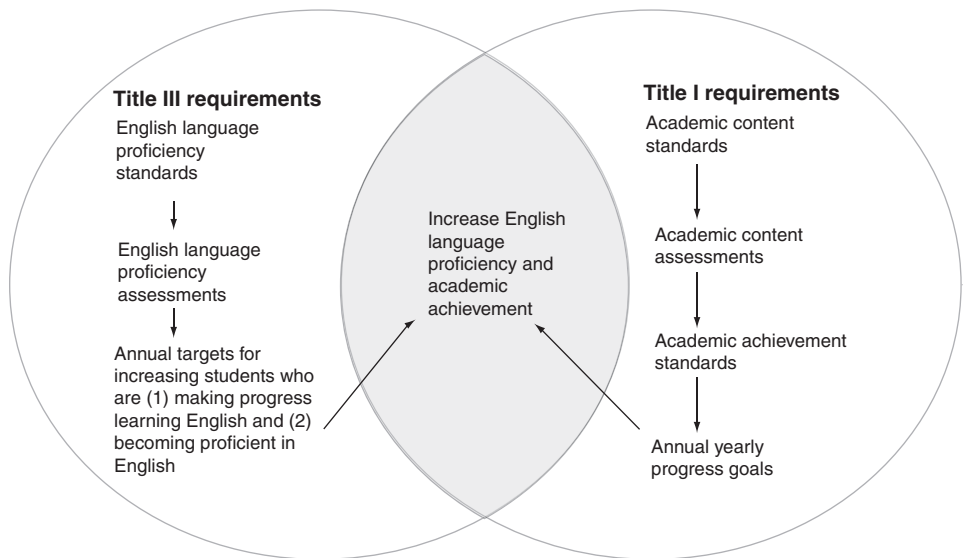
English language proficiency standards	Language arts standards
The student can comprehend reading passages written in familiar or short sentence patterns and verbalize some of the main points of the passage.	The student can independently read and comprehend a grade-level appropriate text and write a short essay describing the main idea of the text.
The student can use acquired knowledge of the English language to learn and understand new vocabulary in context.	The student can apply knowledge of reading strategies to comprehend the text of the next higher level of difficulty.

Source: U.S. Department of Education, “Final Non-Regulatory Guidance on the Title III State Formula Grant Program—Standards, Assessments and Accountability,” February 2003.

Under NCLBA, states, districts, and schools have two sets of responsibilities for students with limited English proficiency. As shown in figure 1, they are responsible for ensuring that these students make progress in learning English under Title III and that they become proficient in language arts and mathematics under Title I. Beginning with the 2004-2005 school year, Education is required to annually review whether states have made adequate yearly progress (as defined by the state) for each of

the student groups and have met their objectives for increasing the number or percentage of students who become proficient in English.

Figure 1: NCLBA’s Requirements for Students with Limited English Proficiency under Title I and Title III



Source: Adapted from U.S. Department of Education, "Final Non-Regulatory Guidance on the Title III State Formula Grant Program—Standards, Assessments and Accountability, February 2003."

Test Development

NCLBA’s emphasis on validity and reliability reflects the fact that these concepts are among the most important in test development. Validity refers to whether the test measures what it is intended to measure. Reliability refers to whether or not a test yields consistent results across time and location and among different sections of the test. A test cannot be considered valid if it is unreliable. The *Standards for Educational and Psychological Testing* provide universally accepted guidance for the development and evaluation of high-quality, psychometrically sound assessments.² They outline specific standards to be considered when assessing individuals with limited English proficiency, including (1) determining when language differences produce threats to the validity and reliability of test results, (2) providing information on how to use and interpret results when tests are used with linguistically diverse individuals,

²These standards are sponsored and published jointly by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education.

and (3) collecting the same evidence to support claims of validity for each linguistic subgroup as was collected for the population as a whole.

Test development begins with determining the purpose of the test and the content to be measured by the test. NCLBA outlines several purposes of statewide assessments, including determining the yearly performance of schools and districts, interpreting individual student academic needs, and tracking the achievement of several groups of students. NCLBA requires that the content of statewide assessments reflects state standards in language arts and mathematics, but the specific skills measured can vary from state to state. For example, a language arts assessment could measure a student's knowledge of vocabulary or ability to write a persuasive essay. Variations in purpose and content affect test design, as well as the analyses necessary to determine validity and reliability.

After determining the purpose and content of the test, developers create test specifications, which delineate the format of the questions and responses, as well as the scoring procedures. Specifications may also indicate additional information, such as the intended difficulty of questions, the student population that will take the test, and the procedures for administering the test. These specifications subsequently guide the development of individual test questions. The quality of the questions is usually ascertained through review by knowledgeable educators and statistical analyses based on a field test of a sample of students—ideally the sample is representative of the overall target student population so the results will reflect how the questions will function when the test is administered to the population. These reviews typically evaluate a question's quality, clarity, lack of ambiguity, and sometimes its sensitivity to gender or cultural issues; they are intended to ensure that differences in student performance are related to differences in student knowledge rather than other factors, such as unnecessarily complex language. Once the quality has been established, developers assemble questions into a test that meets the requirements of the test specifications. Developers often review tests after development to ensure that they continue to produce accurate results.

Education's Responsibilities

Education has responsibility for general oversight of Titles I and III of NCLBA. The department's Office of Elementary and Secondary Education oversees states' implementation of Title I requirements with respect to academic assessments and making adequate progress toward achieving academic proficiency for all students by 2014. Education's Office of English Language Acquisition, Language Enhancement and Academic

Achievement for Limited English Proficient Students oversees states' Title III responsibilities, which include administering annual English language proficiency assessments to students with limited English proficiency and demonstrating student progress in attaining English language proficiency.

Students with Limited English Proficiency Performed below Progress Goals in 2004 in Two-Thirds of States, but States We Studied Are Working to Improve Student Academic Performance

In school year 2003-2004, the percentage of students with limited English proficiency reported by states as scoring proficient on a state's language arts and mathematics tests was lower than the state's annual progress goals (established for all students) in nearly two-thirds of the 48 states for which we obtained data.³ Further, data from state mathematics tests showed that these students were generally achieving lower rates of academic proficiency than the total student population. However, factors other than student academic performance can influence whether a state meets its progress goals, such as which students a state includes in the limited English proficient group and how a state establishes its annual progress goals. Officials in our study states reported using several common approaches, including providing teacher training specific to the needs of limited English proficient students and using data to guide instruction and identify areas for improvement.

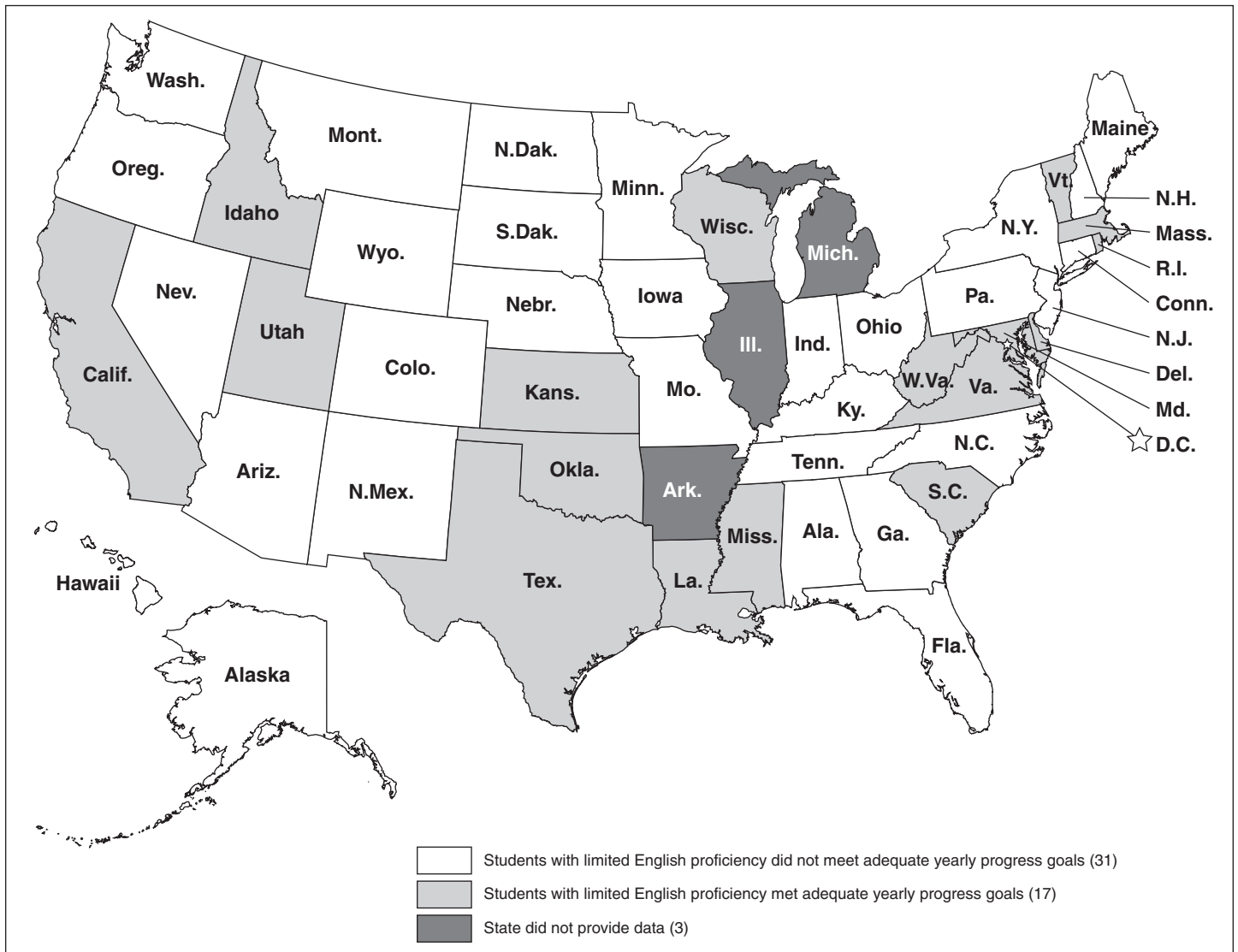
³We included the District of Columbia in our 48-state total.

In Almost Two-Thirds of States, the Percentage of Students with Limited English Proficiency Achieving Proficient Scores Was Below the State's Annual Progress Goals

In nearly two-thirds of the 48 states for which we obtained data, state data showed that the percentage of students with limited English proficiency scoring proficient on language arts and mathematics tests was below the annual progress goal set by the state for school year 2003-2004. Students with limited English proficiency met academic progress goals in language arts and mathematics in 17 states.⁴ In 31 states, state data indicated that these students missed the goals either for language arts or for both language arts and mathematics (see fig. 2). In 21 states, the percentage of proficient students in this group was below both the mathematics and the language arts proficiency goals. See appendix II for information on how adequate yearly progress measures are calculated.

⁴In 7 of the 17 states, students with limited English proficiency met a state's adequate yearly progress goals through NCLBA's safe harbor provision—that is, by decreasing the percentage of students scoring nonproficient by 10 percent or more and showing progress on another academic indicator.

Figure 2: School Year 2003-2004 Comparison of Percentage of Students with Limited English Proficiency Who Achieved Proficient Scores in Language Arts and Mathematics with State-Established Progress Goals



Source: State 2003–2004 report cards available on state Web sites or data provided by state officials.

Notes: We obtained data for 42 states from their state Web sites and contacted state officials in 6 states to obtain these data. Three states did not report data in a format that allowed us to determine whether the percentage of students with limited English proficiency met or exceeded the annual progress goals established by the state.

When states reported proficiency data for different grades or groups of grades, we determined that students with limited English proficiency met a state’s progress goals if the student group met all proficiency and participation goals for all grades reported. An Education official told us that a state could not make adequate yearly progress if it missed one of the progress goals at any grade level.

All of the states on the map where the proficiency percentage for students with limited English proficiency met or exceeded the state's annual progress goal also met NCLBA's participation goals.

We incorporated states' use of confidence intervals and NCLBA's safe harbor provision in determining whether the percentage of students with limited English proficiency achieving proficient scores met or exceeded a state's progress goals. If a state's published data did not explicitly include such information, we contacted state officials to ensure that the state did not meet its progress goals through the use of confidence intervals or through NCLBA's safe harbor provision. In the following seven states, the percentage of students with limited English proficiency was below the state's annual progress goal for language arts or for both language arts and mathematics, but the student group met the state's requirements for progress through the safe harbor provision: Delaware, Idaho, Maryland, Massachusetts, Oklahoma, Rhode Island, and Utah.

We reported 2004–2005 school year data for Oklahoma, New Mexico and Utah because we could not obtain data for the 2003–2004 school year. Data from Iowa, Massachusetts, and Rhode Island are for the 2002–2004 school years.

Rhode Island did not separately report participation rates for students with limited English proficiency. Instead, it reported that all students met the 95 percent participation goal.

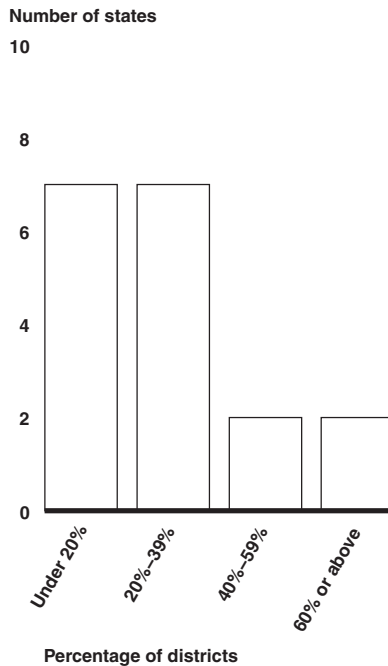
We also obtained additional data from 18 states to determine whether districts were meeting annual progress goals for students with limited English proficiency in school year 2003–2004.⁵ In 14 of the 18 states, however, we found that less than 40 percent of the districts in each state reported separate results for this group of students (see fig. 3).⁶ Districts only have to report progress results for a student group if a minimum number of students are included in the group.⁷ In Nebraska, for example, only 4 percent of districts reported progress goals for students with limited English proficiency. Except for Florida, Hawaii, and Nevada, less than half of the districts in each state reported separate results for this group of students. Even when districts do not have to report on students with limited English proficiency, however, the test scores for these students are included in the state's overall progress measures.

⁵We requested district data from 20 states, and 18 states responded to our request.

⁶The number of districts reporting results for these students may increase in the future, since states were required to assess students in more grades beginning with the 2005–2006 school year. Testing in more grades will likely increase the number of districts meeting the minimum number of limited English proficient students that will be required to separately report proficiency results.

⁷States determine the minimum number of students in a student group, usually between 25 and 45 students.

Figure 3: Percentage of Districts in 18 Selected States Reporting Adequate Yearly Progress Results in School Year 2003-2004 for Students with Limited English Proficiency



Source: GAO analysis of district report cards and district data provided by state officials.

Notes: If a district reported annual progress results for students with limited English proficiency in either language arts proficiency or mathematics proficiency, or both, we considered that the district reported adequate yearly progress results for the student group.

Hawaii has only one school district. Since the state reported separate results for students with limited English proficiency, it has been included as 100 percent of districts reporting separate results for these students.

For those districts that reported results for students with limited English proficiency, district-level data showed that most districts in 13 of the 18 states met their mathematics progress goals for these students. For example, 67 percent of reporting districts in Nebraska and 99 percent of reporting districts in Texas met the state’s goals. In 4 states, less than half of the districts reporting results for these students met the state mathematics progress goals. Specifically, 26 percent of Alaska districts, 33 percent of Nevada districts, 48 percent of Oregon districts, and 48 percent of Florida districts met these goals. (See app. III for results from each of the 18 states.)

In addition to looking at whether students with limited English proficiency met annual progress goals at the state and district level, we also examined

achievement levels on state assessments for this group of students compared with the total student population (which also includes students with limited English proficiency). Looking at mathematics results reported by 49 states to Education, for example, in all but one state, we found that a lower percentage of students with limited English proficiency at the elementary school level achieved proficient scores, compared to the total student population in school year 2003-2004 (see app. IV for the results reported by the 49 states). Twenty-seven states reported that the total student population outperformed students with limited English proficiency by 20 percentage points or more. The differences among groups in the percentage of students achieving proficient scores varied across states. South Dakota, for example, reported a large achievement gap, with 37 percent of limited English proficient students scoring at the proficient level, compared to 78 percent for the entire student population. The gap was less pronounced in Texas, where 75 percent of students with limited English proficiency achieved proficient scores on the mathematics assessment, while 85 percent of the total student population did. In Louisiana, these students performed about the same as the total student population, with 58 percent of limited English proficient students scoring at the proficient level on the elementary mathematics assessment, compared to 57 percent of the total student population.

We also found that, in general, a lower percentage of students with limited English proficiency achieved proficient test scores than other selected student groups (see table 4). All of the 49 states reported that these students achieved lower rates of proficiency than white students.⁸ The performance of limited English proficient students relative to the other student groups varied. In 37 states, for example, economically disadvantaged students outperformed students with limited English proficiency, while students with disabilities outperformed these students in 14 states. In 12 states, all the selected student groups outperformed students with limited English proficiency.

⁸Student groups are not mutually exclusive, with each of the ethnic and racial categories probably including some number of students with limited English proficiency. For example, the results for a student who is both white and limited English proficient would be included in both groups.

Table 4: Percentage of Elementary Students Scoring at the Proficient Level or Above on State Mathematics Assessment for Selected Student Groups, School Year 2003-2004

States	Students with Limited English Proficiency	Students with Disabilities	African-American	Economically Disadvantaged	Hispanic	White
Ala.	53	31	58	62	61	81
Alaska	40	36	50	50	63	77
Ariz. ^a	32	31	46	data not available	44	72
Ark. ^b	49	24	38	53	56	74
Calif.	33	24	28	33	33	61
Colo. ^a	76	61	74	79	79	94
Conn.	47	49	69	61	61	88
Del. ^a	70	47	61	67	74	87
D.C.	34	14	49	48	57	89
Fla.	48	39	44	52	60	74
Ga.	53	46	65	66	67	84
Hawaii ^a	9	6	19	18	16	36
Idaho	62	55	69	76	68	87
Ill. ^a	53	57	54	61	64	89
Ind. ^a	47	40	54	60	60	75
Iowa	49	39	46	62	56	80
Kans.	58	67	61	70	65	84
Ky. ^a	32	29	28	36	38	51
La.	58	30	40	48	62	74
Maine	10	13	15	20	20	33
Md.	39	39	52	51	59	83
Mass.	22	20	20	22	19	49
Mich.	59	42	51	57	58	77
Minn. ^a	38	45	39	52	45	77
Miss.	79	61	69	72	87	91
Mo.	30	24	24	28	29	45
Mont.	15	22	32	33	36	49
Nebr.	73	65	72	79	80	90
Nev. ^a	22	22	27	32	32	57
N.H. ^a	57	61	72	71	67	85
N.J.	47	46	50	54	59	81

States	Students with Limited English Proficiency	Students with Disabilities	African-American	Economically Disadvantaged	Hispanic	White
N.Mex.	36	31	50	47	49	72
N.C.	86	75	88	89	90	≥95
N.Dak.	30	38	45	52	47	68
Ohio	48	38	39	48	51	72
Okla. ^a	41	23	38	46	47	61
Oreg. ^a	61	57	71	73	63	86
Pa. ^a	34	27	30	42	38	70
R.I.	23	34	32	36	31	60
S.C.	16	16	19	22	24	49
S.Dak.	37	48	56	65	62	83
Tex.	75	76	75	80	81	93
Utah	53	43	56	71	52	78
Vt.	67	37	51	58	data not available	73
Va. ^a	84	74	77	79	84	92
Wash.	27	29	38	44	39	66
W.Va.	68	33	53	64	68	70
Wisc.	50	51	45	56	53	80
Wyo.	15	21	25	29	24	42
Number of states where students with limited English proficiency had lower proficiency levels than this group (bolded numbers)		14	28	37	41	49

Source: Consolidated State Performance Reports, 2003-04 school year.

Notes: Bolded numbers indicate when the percentage of students in a group achieving proficient scores is greater than the percentage of students with limited English proficiency achieving proficient scores.

Student groups are not mutually exclusive, so that results for a student who is both Hispanic and economically disadvantaged appear in both groups.

New York and Tennessee did not provide assessment data in their State Consolidated Performance Reports for the 2003-2004 school year.

^aMost states reported assessment data for students in the fourth grade. States marked with this superscript reported on some other grade at the elementary school level, usually either third grade or fifth grade.

^bThe percentages for Arkansas do not include those students with limited English proficiency or students with disabilities who were considered proficient based on the state's portfolio assessment.

Factors beyond Student Performance Influence Progress Measures Reported by States

Factors beyond student performance can influence the number of states, districts, and schools meeting progress goals for students with limited English proficiency. One factor that can affect a state or district's ability to meet progress goals for this student group is the criteria states use to determine which students are counted as limited English proficient. Some states define limited English proficiency so that students may be more likely to stay in the group for a longer time, giving them more of an opportunity to develop the language skills necessary to demonstrate their academic knowledge on state academic assessments administered in English. On the basis of our review of state accountability plans, we found that some states removed students from the group after they have achieved proficiency on the state's English language proficiency assessment, while other states continued to include these students until they met additional academic requirements, such as achieving proficient scores on the state's language arts assessment. A number of states measured adequate yearly progress for students with limited English proficiency by including test scores for students for a set period of time after they were considered proficient in English, following Education's policy announcement in February 2004 allowing such an approach.

How rigorously a state defines the proficient level of academic achievement can also influence the ability of states, districts, and schools to meet their progress goals. States with less rigorous definitions of proficiency are more likely to meet their progress goals for students with limited English proficiency or any other student group than states with more stringent definitions. Comparing the performance of students from different states on a nationally administered assessment suggests that states differ in how rigorously they define proficiency. For example, eighth-grade students in Colorado and Missouri achieved somewhat similar scores on the National Assessment of Educational Progress in mathematics in 2003.⁹ Specifically, 34 percent of Colorado students scored proficient or above on this national assessment compared to 28 percent of Missouri students. On their own state assessments in 2003, however, 67 percent of Colorado students scored proficient or above, compared to

⁹Under NCLBA, states are required to participate in the National Assessment of Educational Progress for reading and mathematics assessments in grades four and eight. The results from these assessments provide a national measure of student achievement and serve as confirmatory evidence about student achievement on state tests.

just 21 percent in Missouri.¹⁰ These results may reflect, among other things, a difference in the level of rigor in the tests administered by these states. However, they may also be due in part to differences in what the national test measures versus what each of the state tests measure.

The likelihood of a state, district, or school meeting its annual progress goals also depends, in part, on the proficiency levels of its students when NCBLA was enacted, as well as how the state sets its annual goals. States vary significantly in the percentage of students scoring at the proficient level on their academic assessments, so that states with lower proficiency levels must, on average, establish larger annual increases in proficiency levels to meet the 2014 goal. Some states planned for large increases every 2 to 3 years, while others set smaller annual increases. States that established smaller annual increases in their initial proficiency goals may be more likely to meet their progress goals at this time, compared with states that set larger annual increases.

The use of statistical procedures, such as confidence intervals, can also affect whether a state, district, or school meets its progress goals. Education officials said that states use such procedures to improve the reliability of determinations about the performance of districts. According to some researchers, such methods may improve the validity of results because they help to account for the effect of small group sizes and year-to-year changes in student populations.¹¹ Most states currently use some type of confidence interval to determine if a state or district has met its progress goals, according to the Center on Education Policy.¹² A confidence interval establishes a range of proficiency levels around a

¹⁰Example taken from Robert L. Linn, "CRESST Policy Brief 8: Fixing the NCLB Accountability System," Policy Brief of the National Center for Research on Evaluation, Standards, and Student Testing, Summer 2005.

¹¹Theodore Coladarci, *Gallup Goes to School: The Importance of Confidence Intervals for Evaluating "Adequate Yearly Progress" in Small Schools*, the Rural School and Community Trust Policy Brief, Oct. 2003, and Thomas J. Kane and Douglas O. Steiger, "Volatility in School Test Scores: Implications for Test-Based Accountability Systems," in Diane Ravitch, ed., *Brookings Papers on Education Policy 2002*, pp. 235-283. Washington, D.C.: Brookings Institution.

¹²Center on Education Policy, "From the Capital to the Classroom: Year 4 of the No Child Left Behind Act," March 2006.

state's annual progress goal.¹³ If the percentage of students with limited English proficiency scoring proficient on a state's academic assessments falls within that range, that group has made the annual progress goal.

States and Districts We Visited Have Taken Steps to Improve Performance of Students with Limited English Proficiency

To help students with limited English proficiency progress academically, state and district officials in our 5 study states reported using somewhat similar strategies, many of which are also considered good practices for all students. Among the key factors cited by state and district officials for their success in working with this group were

- strong professional development focused on effective teaching strategies for students with limited English proficiency;
- school or district leadership that focuses on the needs of these students, such as providing sufficient resources to meet those needs and establishing high academic standards for these students;
- “data driven” decisions, such as using data strategically to identify students who are doing well and those who need more help, to identify effective instructional approaches, or to provide effective professional development; and
- efforts to work with parents to support the academic progress of their children.

These approaches are similar to those used by “blue ribbon” schools—schools identified by Education as working successfully with all students to achieve strong academic outcomes. The qualities shared by these blue ribbon schools include professional development related to classroom instruction, strong school leadership and a vision that emphasizes high academic expectations and academic success for all students, using data to target instructional approaches, and parental involvement. While many blue ribbon schools have a high percentage of disadvantaged students, including those with limited English proficiency, their common approaches help them achieve student outcomes that place them among the top 10 percent of all schools in the state or that demonstrate dramatic improvement.

¹³When using confidence intervals, upper and lower limits around a district's or state's percentage of proficient students are calculated, creating a range of values within which there is “confidence” the true value lies. For example, instead of saying that 72 percent of students scored at the proficient level or above on a test, a confidence interval may show that percentage to be between 66 and 78 percent, with 95 percent confidence.

Officials in all 5 of our study states stressed the importance of providing teachers with the training they need to work effectively with students with limited English proficiency. For example, state officials in North Carolina told us that they are developing a statewide professional development program to train mainstream teachers to present academic content material so that it is more understandable to students with limited English proficiency and to incorporate language development while teaching subjects such as mathematics and science. In one rural North Carolina school district where students with limited English proficiency have only recently become a large presence, district officials commented that this kind of professional development has helped teachers become more comfortable with these students and given them useful strategies to work more effectively with them.

In 4 of our study states, officials emphasized the need for strong school or district leadership that focuses on the needs of students with limited English proficiency. For example, officials in a California school district with a high percentage of students with limited English proficiency told us that these students are a district priority and that significant resources are devoted to programs for them. The district administration has instilled the attitude that students with limited English proficiency can meet high expectations and are the responsibility of all teachers. To help maintain the focus on these students, the district has created an English language development progress profile to help teachers track the progress of each student in acquiring English and meeting the state's English language development standards.

In addition, officials in 4 of our study states attributed their success in working with students with limited English proficiency to using data strategically, for example, to identify effective practices and guide instruction. At one California school we visited, officials reviewed test scores to identify areas needing improvement for different classes and different student groups and to identify effective practices. In addition, they reviewed test data for each student to identify areas of weakness. If test data showed that a student was having trouble with vocabulary, the teacher would work in class to build the student's vocabulary. Similarly, officials in a New York school reported that they followed student test scores over 3 years to set goals for different student groups and identify areas in need of improvement.

Officials in 3 states we visited also cited the importance of involving parents of students with limited English proficiency in their children's education. In Nebraska, for example, a technical assistance agency

implemented a family literacy program to help parents and their children improve their English, and also to involve parents in their children's education. The program showed parents how they can help children with their homework and the importance of reading to their children in their native language to develop their basic language skills. At a New York middle school, officials told us that they use a parent coordinator to establish better communication with families, learn about issues at home that can affect the student's academic performance, and help families obtain support services, if needed.

Selected States Considered Language Issues When Developing Academic Assessments, but Validity and Reliability Concerns Remain

For academic assessments in language arts and mathematics, officials in the 5 states we studied reported that they have taken some steps, such as reviewing test items to eliminate unnecessarily complex language, to address the specific challenges associated with assessing students with limited English proficiency. However, Education recently reviewed the assessment documentation of 38 states and noted some concerns related to using these assessments for students with limited English proficiency. Our group of experts also indicated that states are generally not taking the appropriate set of comprehensive steps to create assessments that produce valid and reliable results for students with limited English proficiency. To increase the validity and reliability of assessment results for this population, most states offered accommodations, such as providing extra time to complete the assessment and offering native language assessments. However, offering accommodations may or may not improve the validity of test results, as research on the appropriate use of accommodations for these students is lacking. In addition, native language assessments are not appropriate for all students with limited English proficiency and are difficult and expensive to develop.

States Reported Efforts to Improve Validity of Assessment Results for Students with Limited English Proficiency

Officials in the 5 states we studied reported taking some steps to address the specific challenges associated with assessing students with limited English proficiency in language arts and mathematics. Officials in 4 of these states reported following generally accepted test development procedures when developing their academic assessments, while a Nebraska official reported that the state expects districts to follow such procedures when developing their tests. Test development involves a structured process with specific steps; however, additional steps and special attention to language issues are required when developing a test that includes students with limited English proficiency to ensure that the results are valid and reliable for these students. As the *Standards for Educational and Psychological Testing* notes, for example, the test

instructions or the response format may need to be modified to ensure that the test provides valid information about the skills of students with limited English proficiency.

Officials in 2 states and at several testing companies mentioned that they have been focusing more on the needs of these students in recent years. Officials in California, New York, North Carolina, and Texas told us that they try to implement the principles of universal design, which support making assessments accessible to the widest possible range of students. This is done by ensuring, among other things, that instructions, forms, and questions are clear and not more linguistically complex than necessary. In addition, officials in all 5 states we studied told us they included students with limited English proficiency in the field testing of assessments. North Carolina officials reported that they oversample for students with limited English proficiency to ensure that these students are adequately represented in the field tests.

Another step officials in some states reported taking is assembling panels or committees to review test items for bias and testing data for bias related to a student's English proficient status. For example, Texas and North Carolina officials reported creating review committees to ensure that test items are accessible to students with limited English proficiency. Specifically, when developing mathematics items, these states try to make the language as clear as possible to ensure that the item is measuring primarily mathematical concepts and to minimize the extent to which it is measuring language proficiency. A mathematics word problem involving subtraction, for example, might refer to fish rather than barracuda. Officials in 4 of our study states told us they used a statistical approach to evaluate test items for bias against specific student groups, and three of these reported using it to detect bias related to students with limited English proficiency. However, this type of analysis can only be used when a relatively large number of students in the specific group is taking the test. Members of our expert group recommended the use of this technique for states with a large enough number of students with limited English proficiency; however, one member noted that this technique may not be appropriate if a state's population of students with limited English proficiency is diverse but is treated as homogenous in the analyses.

Some of our study states also reported including experts on limited English proficiency or English as a second language (ESL) issues in the development and review of test items, although only 1 reported involving them in all aspects of test development. In North Carolina, for example, officials told us that ESL teachers and supervisors are involved in

reviewing all aspects of the test development process, including item writing, field testing, and operational testing. Some state officials also told us that they included education staff involved with students with limited English proficiency in the development of assessments.

Both Education's Peer Reviews and Our Group of Experts Raised Concerns Regarding State Efforts to Ensure Valid and Reliable Results

Education's recent NCLBA peer reviews of 38 states¹⁴ found that 25 did not provide sufficient evidence on the validity or reliability of results for students with limited English proficiency, although states have been required to include these students in their assessments since 1994.¹⁵ For example, peer reviewers found that Alabama's documentation did not include sufficient evidence on the selection process for committee members to review test items for bias, noting that no evidence was provided on whether representatives for students with limited English proficiency were included. In Idaho, peer reviewers commented that the state did not report reliability data for student groups, including students with limited English proficiency. See table 5 for further examples.

¹⁴As of July 2006, Education had conducted peer reviews of 50 states and the District of Columbia. However, detailed peer review notes were available from only 38 states at the time of our review.

¹⁵States were given several years to meet the requirements of the Improving America's School Act of 1994.

Table 5: Examples of Issues Relating to Assessing Students with Limited English Proficiency Raised in Education’s Peer Review Reports

Validity	Reliability	Accommodations
The state’s item development and review procedures are inadequate to ensure that the assessments do not reflect unfair irrelevant constructs. For example, there is no empirical evidence that the item review process is successful in removing barriers due to overly complex language. Further, a statistical process to determine bias is evaluated only for gender and race/ethnicity; the state should consider using it to evaluate geographic and demographic diversity, such as students with limited English proficiency.	The state did not provide any reliability data for each reported subpopulation.	The state provides a reasonable list of accommodations, but does not distinguish among those that are allowable for students with disabilities, and those which are allowable for students with limited English proficiency. The state may wish to provide separate lists of accommodations to support the selection of appropriate accommodations that are aligned with instructional approaches for individual students. Further, although it appears that the state has a system in place for monitoring the selection of accommodations for students with disabilities, it does not for students with limited English proficiency.
The state should conduct a bias review, especially for the common items and the alternate assessments for students with limited English proficiency and students with disabilities.	The state does not routinely report subgroup reliability data.	The state did not provide evidence to support the assertion that accommodations for students with limited English proficiency allow for valid inferences about these students’ knowledge and skills. It does not appear that the state monitors availability of accommodations during test administration. The use of accommodations should be validated on the state’s own student population.

Source: GAO review of Education’s peer review notes for 38 states.

Note: This table includes examples from the categories used in Education’s peer review process that determine if a state’s assessments are valid, reliable, fair and accessible, and use appropriate accommodations.

Our group of experts indicated that states are generally not taking the appropriate set of comprehensive steps to create assessments that produce valid and reliable results for students with limited English proficiency and identified essential steps that should be taken. The group noted that no state has implemented an assessment program for students with limited English proficiency that is consistent with the *Standards for Educational and Psychological Testing* and other technical standards. Specifically, the group said that students with limited English proficiency are not defined consistently within and across states, which is a crucial first step to ensuring the reliability of test results. A reliable test should produce consistent results, so that students achieve similar scores if tested repeatedly. If the language proficiency levels of students with limited English proficiency are classified inconsistently, an assessment may produce results that appear inconsistent because of the variable classifications rather than actual differences in skill levels. One expert noted, for example, that some studies have shown that a student’s language proficiency plays a small role in determining whether a student is classified as limited English proficient. Inconsistency in defining these

students may be due to variation in how school districts apply state definitions. For example, according to a 2005 study on students with limited English proficiency in California, state board of education guidelines suggest that districts consider a student's performance on the state's English language proficiency assessment and on the state's language arts test, a teacher evaluation of the student's academic performance, and parental recommendations when determining if a student should or should not continue to be considered limited English proficient. However, the study noted that districts interpreted and applied these factors differently.¹⁶ Further, it appears that many state assessment programs do not conduct separate analyses for different groups of limited English proficient students. Our group of experts indicated that the reliability of a test may be different for heterogeneous groups of students with limited English proficiency, such as students who are literate in their native language and those who are not.

Our group of experts also noted that states are not always explicit about whether an assessment is attempting to measure skills only (such as mathematics) or mathematics skills as expressed in English. According to the group, a fundamental issue affecting the validity of a test is the definition of what is being measured. Members of the group emphasized that approaches to ensure valid test results should vary based on which of these is being measured. For example, North Carolina officials stated that the state did not offer native language assessments because the state has explicitly chosen to measure student knowledge in English.

The expert group emphasized that determining the validity and reliability of academic assessments for students with limited English proficiency is complicated and requires a comprehensive collection of evidence rather than a single analysis or review. As one expert noted, "you can't just do one thing and assume things are valid." In addition, the appropriate combination of analyses will vary from state to state, depending on the characteristics of the student population and the type of assessment. For example, because reliability of test results can vary based on a student's English proficiency status or a student's native language, states with more diverse groups of limited English proficient students may need to conduct additional analyses to ensure sufficient reliability. The group indicated that states are not universally using all the appropriate analyses to

¹⁶Christopher Jepsen and Shelley de Alth, "English Learners in California Schools," (San Francisco, California: Public Policy Institute of California, 2005).

evaluate the validity and reliability of test results for students with limited English proficiency. Instead, our experts noted that states vary in terms of the particular techniques they use for this purpose, and in the extent to which they collect valid background data. Members indicated that some states may need assistance to conduct appropriate analyses that will offer useful information about the validity of their academic assessments for these students.

Finally, our group of experts indicated that reducing language complexity is essential to developing valid assessments for these students, but expressed concern that some states and test developers do not have a strong understanding of universal design principles or how to use them to develop assessments from the beginning to eliminate language that is not relevant to measuring a student's knowledge of, for example, mathematics. Members believed that some states may need more information on how to implement these principles to develop assessments that produce valid results for students with limited English proficiency.

Accommodations Can Increase Validity of Assessment Results, but Research on Appropriate Use Is Limited

The majority of states offered some accommodations to try to increase the validity and reliability of assessment results for students with limited English proficiency. These accommodations are intended to permit students with limited English proficiency to demonstrate their academic knowledge, despite their limited language ability. Our review of state Web sites found available documentation on accommodations for 42 states. The number of accommodations offered varied considerably among states. One state, for example, offered students with limited English proficiency the use of a bilingual dictionary and a side-by-side English-Spanish version of its grade 10 mathematics test. Another state listed over 40 acceptable accommodations, including clarifying test directions in English or the student's native language, offering extra time, and providing responses (written or oral) in the student's native language.

Our review found that the most common accommodations offered by these states were allowing the use of a bilingual dictionary and reading test items aloud in English (see table 6). In addition, they offered other accommodations intended to create a less distracting environment for students, such as administering the assessment to the student in a small group or individually. Some states also gave students with limited English proficiency extra time to complete a test to account for their slower reading speed and information processing time in English. The 5 states we studied varied in how they established and offered accommodations to students. For example, Texas officials reported working with its limited

English proficiency focus group to develop a list of allowable accommodations, which may be offered on a test when they are routinely used by students in their classrooms. In addition, each school district has a committee to select particular accommodations based on the needs of individual students. California officials told us the state provides guidance to districts on the appropriate use of accommodations. However, they said that districts might not provide approved accommodations because of high administrator turnover.

Table 6: Most Frequently Cited Accommodations in 42 States

Accommodation	Number of states
Bilingual dictionary	32
Reading items aloud in English	32
Small group administration	29
Extra time	27
Individual administration	27
Separate location	25
Extra breaks	25
Directions in student's native language	24

Source: GAO review of state documentation.

According to our expert group and our review of the relevant literature, research is lacking on what specific accommodations are appropriate for students with limited English proficiency, as well as their effectiveness in improving the validity of assessment results. A 2004 review of state policies found that few studies focus on accommodations intended to address the linguistic needs of students with limited English proficiency or on how accommodations affect the performance of students with limited English proficiency.¹⁷ In contrast, significantly more research has been conducted on accommodations for students with disabilities, much of it funded by Education. Because of this research disparity, our group of experts reported that some states offer accommodations to students with limited English proficiency based on those they offer to students with disabilities, without determining their appropriateness for individual students. Our experts noted the importance of considering individual student characteristics to ensure that an accommodation appropriately

¹⁷Charlene Rivera and Eric Collum. *An Analysis of State Assessment Policies Addressing the Accommodation of English Language Learners*. The George Washington University Center for Equity and Excellence in Education, Arlington, Virginia: (January 2004).

addresses the needs of the student. Other researchers have raised similar issues about the use of accommodations by states.

Education's peer reviews of state academic assessments identified issues related to accommodations for students with limited English proficiency in all 38 states reviewed. For example, the reviewers noted that South Dakota does not clearly indicate whether students with limited English proficiency were provided accommodations that they do not regularly use in the classroom. If an accommodation is not used regularly in the classroom, it may not improve the validity of test results because the student may not be comfortable with a new procedure. In addition, they noted that South Dakota does not appear to be monitoring the use of accommodations and suggested that the state study accommodations to ensure that they are instructionally appropriate and that they improve the validity and reliability of the results. In Texas, the reviewers noted that the state needs to provide information regarding the quality and consistency of accommodations for students with limited English proficiency—specifically whether the state routinely monitors the use of accommodations for these students. In North Carolina, they noted a lack of evidence that the state has performed research on accommodations. Although conducting such research could provide useful information on the validity of accommodated tests, having each state individually study accommodations could be financially burdensome for them. While research on accommodations for this population would be useful, it does not have to be conducted directly by states to be applicable to a state's student population. Further, such research could involve short-term studies, rather than large-scale, longitudinal efforts.

Native Language and Alternate Assessments May Improve the Validity of Results but Are Challenging to Implement

In our survey, 16 states reported that they offered statewide native language assessments in language arts or mathematics in some grades for certain students with limited English proficiency in the 2004-2005 school year. For example, New York translated its statewide mathematics assessments into Spanish, Chinese, Russian, Korean, and Haitian-Creole. In addition, 3 states were developing or planning to develop a native language assessment, and several states allowed school districts to translate state assessments or offer their own native language assessments. Our group of experts told us that this type of assessment is difficult and costly to develop. An assessment provided in a student's native language is intended to remove language barriers students face in demonstrating their content knowledge and thereby improve the validity of test results. Of the 16 states that offered statewide native language assessments, 4 were able to provide complete data on the number of

students taking native language assessments. These data indicated that relatively few students took these assessments.

Our group of experts and some state officials also described the challenges of developing native language assessments that produce valid results. Members of our expert group and other experts told us that native language assessments are generally an effective accommodation only for students in specific circumstances, such as students who are instructed in their native language or are literate in their native language. In addition, our experts emphasized that developing valid native language assessments is challenging, time-consuming, and expensive. Development of a valid native language assessment involves more than a simple translation of the original test; in most situations, a process of test development and validation similar to that of the nontranslated test is recommended to ensure the validity of the test. In addition, the administration of native language assessments may not be practicable, for example, when only a small percentage of limited English students in the state speak a particular language or when a state's student population has many languages.

Thirteen states offered statewide alternate assessments (such as reviewing a student's classroom work portfolio) in 2005 for certain students with limited English proficiency, based on our review of accountability plans for all states and the District of Columbia as of March 2006. We also found that 4 additional states allowed school districts to offer alternate assessments, while 7 states and the District of Columbia planned to offer alternate assessments. An official in Wisconsin told us that the state administers an alternate assessment because developing a native language assessment for its relatively small Spanish-speaking population would be impractical and the state does not have bilingual programs in the second most common language, Hmong (a language that is native to Southeast Asia). However, our group of experts noted that alternate assessments are difficult and expensive to develop, and may not be feasible because of the amount of time required for such an assessment. Members of the group also expressed concern about the extent to which these assessments are objective and comparable and can be aggregated with regular assessments. See figure 4 for information on which states offered native language or alternate assessments for students with limited English proficiency.

Most States Implemented New English Language Proficiency Assessments but Faced Challenges Establishing Their Validity and Reliability

With respect to English language proficiency assessments, many states implemented new tests to address NCLBA requirements, and are working to align them with newly required state English language proficiency standards. State and consortia officials reported that states are using assessments or test items developed by state consortia, customized assessments developed by testing companies, state-developed assessments, and off-the-shelf assessments. While a few states already had the required English language proficiency assessments in place, many states are implementing them for the first time in spring 2006; as a result, evidence on their validity and reliability may not be fully developed.

States Are Working with Consortia and Test Developers and Individually to Develop New English Language Proficiency Assessments

Many states implemented new English language proficiency assessments for the 2005-2006 school year to meet Education's requirement for states to administer English language proficiency tests that meet NCLBA requirements by the spring of 2006.¹⁸ These assessments must allow states to track student progress in learning English; in addition, Education requires that these assessments be aligned to a state's English language proficiency standards. According to Education and test development officials, prior to NCLBA, most states used off-the-shelf English language proficiency assessments to determine the placement of students in language instruction programs, but these assessments did not have to be aligned with standards. Education officials said that because many states did not have tests that met NCLBA requirements, the agency funded four state consortia to develop new assessments that were to be aligned with state standards and measure student progress. Officials in some states told us they have chosen to use these consortium-developed tests, while officials in other states reported developing their own tests or continuing to use off-the-shelf tests. Some states had only recently determined what test they are going to administer this year, while others may administer a new test in the 2006-2007 school year. Education officials noted that states' decisions on these tests have been in flux during this transition year.

In the 2005-2006 school year, 22 states used assessments or test items developed by one of four state consortia, making this the most common approach taken by states to develop new English language proficiency

¹⁸Education officials told us that the agency has approved an extension of this deadline for 1 state and is currently considering extension requests from 2 other states.

assessments. Each of the four consortia varied somewhat in its development approach.¹⁹ For example, officials in two consortia reported that they examined all their member states' English language proficiency standards and reached consensus on core standards for use on the English language proficiency assessments. They also planned to continue working with member states in implementing their assessments. For example, one consortium plans to provide ongoing professional development to help educators understand the consortium's standards. In contrast, officials in the other two consortia reported that the consortia disbanded after developing their assessments. One state official told us that the state hired a contractor to customize the consortium-developed assessment to more closely align with state standards. In addition, officials in other states, such as New Mexico, told us they are using a combination of consortium-developed test items, along with items developed by another test developer.

Fifteen states participated in one of the consortia, but officials in these states told us they chose not to use the assessments developed by the consortia in the 2005-2006 school year for a variety of reasons, including lack of alignment with state standards, the length of the assessment, and the cost of implementation. For example, Kentucky chose not to use the consortium assessment because of cost effectiveness concerns and lack of alignment with state standards. Another state decided not to use the consortium-developed assessment, as officials were concerned about its cumbersome nature and associated cost. Officials in some states told us they plan to use consortium-developed assessments in the future. For example, Florida officials reported that the state will use a consortium assessment in the 2006-2007 school year. Appendix V shows the states that participated in the consortia and which used consortia-developed assessments in the 2005-2006 school year.

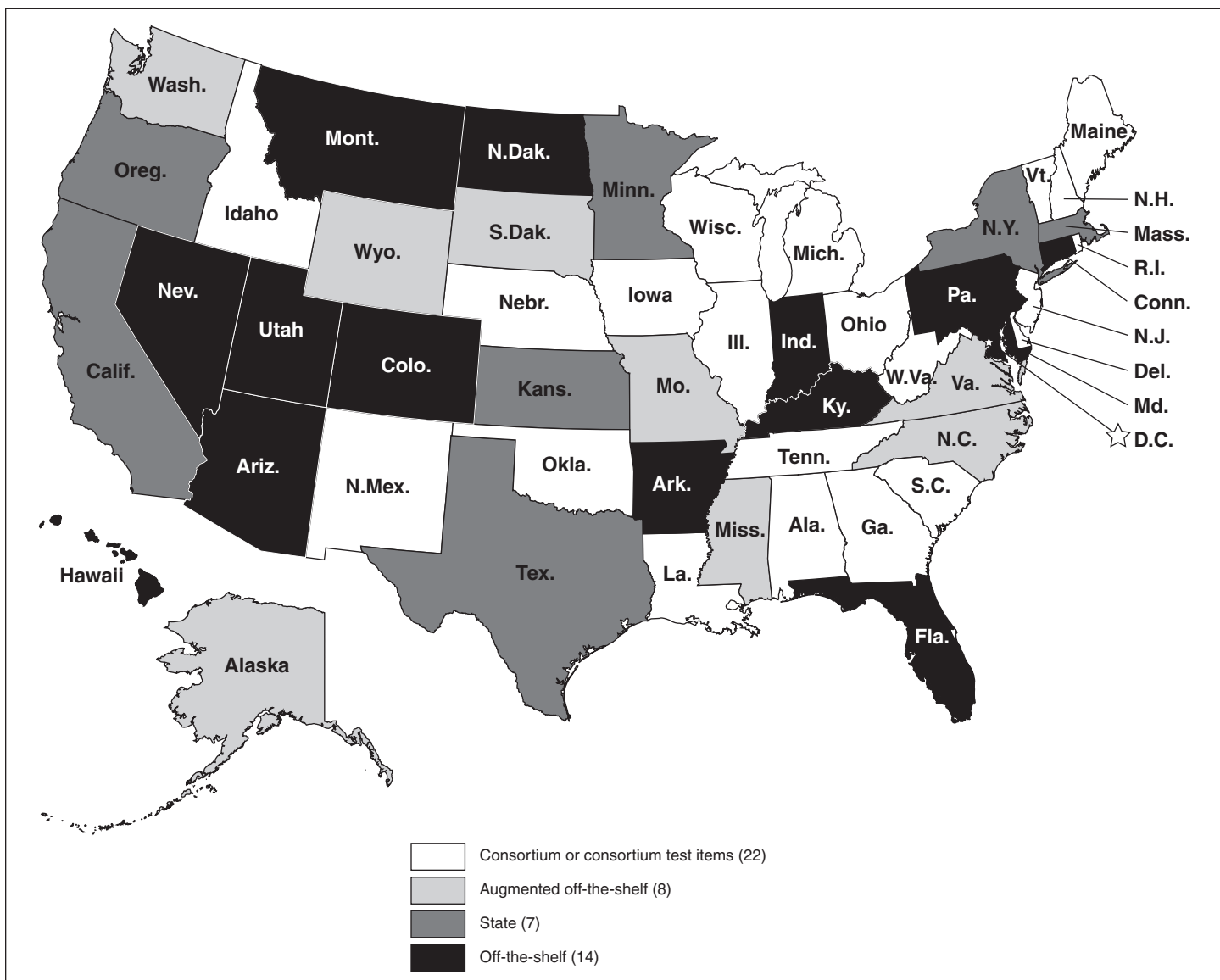
Officials in states that did not use consortia assessments told us that they used other approaches to develop their English language proficiency assessments. Eight states worked with test developers to augment off-the-shelf English language proficiency assessments to incorporate state standards. For example, Mississippi, South Dakota, and Wyoming are using versions of an English language proficiency assessment that has

¹⁹The four consortia include the World-Class Instructional Design and Assessment (WIDA) Consortium, State Collaborative on Assessment and Student Standards (SCASS) Consortium, Mountain West Assessment Consortium (MWAC), and Pennsylvania Enhanced Assessment Grant (PA EAG) Consortium.

been augmented to align to their respective state standards. Officials in 14 states indicated that they are administering off-the-shelf assessments. These officials indicated varying degrees of alignment between the off-the-shelf tests being used and their state's English language proficiency standards; in 11 of these states, the assessment has not been fully aligned with state standards.²⁰ Seven states, including Texas, Minnesota, and Kansas, created their own English language proficiency assessments. Officials in these states said they typically worked with a test developer or research organization to create the assessments. See figure 5 and appendix VI for more detailed information on the English language proficiency assessments used by each state.

²⁰ Although these assessments are not fully aligned to state standards, Education officials noted that they have not yet provided states guidance on what level of alignment the agency expects. Thus, some of these may ultimately meet Education's requirements.

Figure 5: Type of English Language Proficiency Assessment Administered in 2005–2006 School Year



Source: Interviews with consortia and state officials.

Some officials in our 5 study states and 28 additional states we contacted to determine what English language proficiency assessment they planned to use in 2006 pointed to some challenges involving their English language proficiency assessments. Some of these state officials expressed concerns about using both their old and new English language proficiency assessments to measure student progress in learning English. NCLBA

required states to begin tracking student progress in the 2002–2003 school year, before most states had implemented their new English language proficiency assessments. In May 2006, Education officials told us that states must rely on baseline results from their old tests and determine how results from their old tests relate to results from their new tests in order to track student progress since 2003, as required by NCLBA. They noted that states may change their English language proficiency goals based on results from their new assessments, but they cannot change the initial baseline established with their old test. In its technical comments on this report, Education noted that it allows states to make such determinations in a variety of ways, as long as annual progress is reported. Officials in some states want to rely solely on data from their new tests to track student progress. They stated that, unlike their old tests, their new tests provide more accurate data on student progress because they are aligned to their English language proficiency standards and were designed to measure student progress. Officials from other states questioned the usefulness of conducting studies to determine the relationship between their old and new tests, especially in states that had previously used multiple English language proficiency assessments.

Officials in a few of our study states also expressed concern about the appropriateness of NCLBA’s requirement to assess students with limited English proficiency in kindergarten and the first and second grades. For example, Texas officials told us traditional tests do not produce good test results for students this young in part because of their limited attention spans. In addition, officials in Texas and North Carolina noted that English proficient students in these grades are not required to be assessed in the same way.

Many States Are Still in the Process of Establishing the Validity and Reliability of English Language Proficiency Assessments

Officials in our study states and test developers we interviewed reported that they commonly apply generally accepted test development procedures in the development of English language proficiency assessments, but some are still in the process of documenting the validity and reliability of these assessments. For example, some evidence needed to confirm the validity and reliability of the test can be obtained only after the assessment has been fully administered. One consortium contracted with a research organization to assess the validity and reliability testing of its English language proficiency assessment. According to a consortium official, the research organization performed all of the standard steps that are taken to ensure high-quality assessments. These included piloting and field testing the assessment and conducting statistical modeling. An official from another consortium said that its test vendor is conducting

basic psychometric research and analyzing field test data for evidence of reliability. California officials noted that the process for developing and ensuring the validity and reliability of its English language proficiency assessment is similar to that used for its state academic assessments.

Although states have taken steps toward determining validity, documenting the validity and reliability of a new assessment is an ongoing process. A 2005 review of the documentation of 17 English language proficiency assessments used by 33 states in the 2005-2006 school year found that the evidence presented on validity and reliability was generally insufficient.²¹ The report, which was funded by Education, reviewed documentation for consortium-developed assessments, off-the-shelf assessments, and custom-developed assessments for evidence of validity, reliability, and freedom from test bias, among other things. It found that the technical adequacy of English language proficiency assessments is undeveloped compared to the adequacy of assessments for general education. The study noted that none of the assessments contained “sufficient technical evidence to support the high-stakes accountability information and conclusions of student readiness they are meant to provide.”

In addition, many states are in the process of aligning these assessments to state English language proficiency standards, which in turn must be aligned to state content standards. These steps are needed to comply with NCLBA requirements. Alignment, which refers to the degree to which an assessment’s items measure the content they are intended to measure, is critical in assuring the validity of an assessment. Officials in some states²² have expressed uncertainty about how to align their English language proficiency test with their standards for academic subjects, such as mathematics and science.²³ Officials in 2 states told us that their English

²¹Stanley Rabinowitz and Edynn Sato, “Evidence-Based Plan: Technical Adequacy of Assessments for Alternate Student Populations: A Technical Review of High-Stakes Assessments for English Language Learners,” WestEd (December 2005).

²²The states providing these comments represent more than our 5 case study states. We also contacted officials in 28 additional states to determine what English language proficiency assessment they planned to use in 2006.

²³Education does not require state English language proficiency tests to be aligned to state academic standards. However, states’ English language proficiency tests and academic standards are connected, in that Education requires that state’s English language proficiency tests be aligned to state English language proficiency standards and NCLBA requires that state English language proficiency standards be aligned with state academic standards.

language proficiency assessments are aligned to state language arts standards but are not aligned to state mathematics standards, meaning that the assessment may not measure the language needed to succeed in a mathematics class. Findings from Education's Title III monitoring reviews of 13 states indicated that 8 states had not yet fully completed alignment; of these, 5 had not yet linked their English language proficiency and academic content standards, while 5 had not yet aligned their English language proficiency assessments with their English language proficiency standards.²⁴

Education Has Provided Assistance, but States Reported Need for Additional Guidance and Flexibility

Education has offered states a variety of technical assistance to help them appropriately assess students with limited English proficiency, such as providing training and expert reviews of their assessment systems, as well as flexibility in assessing these students. However, Education has issued little written guidance on how states are expected to assess and track the English proficiency of these students, leaving state officials unclear about Education's expectations. To support states' efforts to incorporate these students into their accountability systems, Education has offered states some flexibilities in how they track progress goals for these students. However, many of the state and district officials we interviewed told us that the current flexibilities do not fully account for some characteristics of certain students in this student group, such as their lack of previous schooling. These officials indicated that additional flexibility is needed to ensure that the federal progress measures accurately track the academic progress of these students.

Education Has Provided a Variety of Support on Assessment Issues but Little Written Guidance on Assessing Students with Limited English Proficiency

Education offers support in a variety of ways to help states meet NCLBA's requirements for assessing students with limited English proficiency for both their language proficiency and their academic knowledge. Some of these efforts focus specifically on students with limited English proficiency, while others, such as the Title I monitoring visits, focus on all student groups and on broader compliance issues but review some assessment issues related to students with limited English proficiency as part of their broader purposes. The agency's primary technical assistance efforts have included the following:

²⁴We reviewed reports from Title III monitoring visits of 13 states conducted between April and October 2005 that were available from Education as of March 31, 2006.

-
- *Title I peer reviews of states' academic standards and assessment systems:* Education is currently conducting peer reviews of the academic assessments that states use in measuring adequate yearly progress. During these reviews, three independent experts review evidence provided by the state about the validity and reliability of these assessments (including whether the results are valid and reliable for students with limited English proficiency) and make recommendations to Education about whether the state's assessment system is technically sufficient and meets all legal requirements. Education shares information from the peer review to help states address issues identified during the review. Education has imposed a deadline requiring that states receive peer review approval by June 30, 2006, but only 10 states have had their assessment systems fully approved by Education as of that date.²⁵
 - *Title III monitoring visits:* Education began conducting site visits to review state compliance with Title III requirements in 2005 and has visited 15 states. Education officials reported that they plan to visit 11 more states in 2006. As part of these visits, the agency reviews the state's progress in developing English language proficiency assessments that meet NCLBA requirements.
 - *Comprehensive centers:* Education has contracted with 16 regional comprehensive centers to build state capacity to help districts that are not meeting their adequate yearly progress goals. The grants for these centers were awarded in September 2005, and the centers provide a broad range of assistance, focusing on the specific needs of individual states. At least 3 of these centers plan to assist individual states in developing appropriate goals for student progress in learning English. In 2005, Education also funded an assessment and accountability comprehensive center, which provides technical assistance to the regional comprehensive centers on issues related to the assessment of students, including those with limited English proficiency.
 - *Ongoing technical assistance for English language proficiency assessments:* Education has provided information and ongoing technical assistance to states using a variety of tools and has focused specifically on the development of the English language proficiency standards and assessments required by NCLBA. These include:

²⁵Education has sent letters to the remaining states outlining the issues that need to be resolved in order for their assessment systems to be approved.

-
- a semiannual review of reports states submit to Education and phone calls to state officials focused on state progress in developing their English language proficiency assessments;
 - on-site technical assistance to states regarding their English language proficiency assessments;
 - an annual conference focused on students with limited English proficiency that includes sessions on assessment issues, such as aligning English language proficiency and academic content standards;
 - videoconference training sessions for state officials on developing English language proficiency assessments;
 - providing guidance on issues related to students with limited English proficiency on its Web site;
 - distributing information through an electronic bulletin board and a weekly electronic newsletter focused on students with limited English proficiency;
 - disseminating information through the National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs;
 - semiannual meetings and training sessions with state Title III directors; and
 - responding to questions from individual states as needed.
- *Enhanced Assessment Grants:* Since 2003, Education has awarded these grants, authorized by NCLBA, to support state activities designed to improve the validity and reliability of state assessments. According to an Education official, most of the grants up to now have funded the English language proficiency consortia, although some grants have been used to conduct research on accommodations. For grants to be awarded in 2006, Education will give preference to projects involving accommodations and alternate assessments intended to increase the validity of assessments for students with limited English proficiency and students with disabilities.
 - *Title I monitoring visits:* As part of its monitoring visits to review state compliance with Title I requirements, Education reviews some aspects of the academic assessments administered by states, but in less detail than during its peer reviews. During these visits, for example, states may receive some feedback on how the state administers academic assessments to students with limited English proficiency and the appropriateness of accommodations offered to these students. Education staff also reported that they respond to questions about Title I requirements from individual states as needed.

While providing states with a broad range of technical assistance and guidance through informal channels, Education has issued little written

guidance on developing English language proficiency assessments that meet NCLBA's requirements and on tracking the progress of students in acquiring English. Education issued some limited nonregulatory guidance on NCLBA's basic requirements for English language proficiency standards and assessments in February 2003. However, officials in about one-third of the 33 states we visited or directly contacted expressed uncertainty about implementing these requirements. They told us that they would like more specific guidance from Education to help them develop tests that meet NCLBA requirements, generally focusing on two issues. First, some officials said they were unsure about how to align English language proficiency standards with content standards for language arts, mathematics, and science, as required by NCLBA. An official in 1 state said the state needed specific guidance on what Education wants from these assessments, such as how to integrate content vocabulary on the English language proficiency assessment without creating an excessively long test. In another state, officials explained that the state was developing its English language proficiency test by using an off-the-shelf test and incorporating additional items to align the test with the state's English language proficiency and academic standards. However, the state discovered that it had not correctly augmented the test and will consequently have to revise the test. Officials in this state noted that they have had to develop this test without a lot of guidance from Education.

Second, some officials reported that they did not know how to use the different scores from their old and new English language proficiency assessments to track student progress. For example, an official in 1 state said that she would like guidance from Education on how to measure student progress in English language proficiency using different tests over time. Another official was unsure if Education required a formal study to correlate the results from their old and new English language proficiency assessments, noting that more specific guidance would help them better understand Education's requirements. Without guidance and specific examples on both of these issues, some of these officials were concerned that they will spend time and resources developing an assessment that may not meet Education's requirements.

Education officials told us that they are currently developing additional nonregulatory guidance on these issues, but it has not been finalized. They also pointed out that they have provided extensive technical assistance on developing English language proficiency standards and assessments, and have clearly explained the requirements to state officials at different meetings on multiple occasions. An Education official acknowledged that states were looking for more guidance on the degree of alignment required

between their English language proficiency assessments and standards, noting that Education is still considering the issue. She stated that the issue would be addressed in the guidance it plans to issue in the future.

With respect to academic content assessments, our group of experts reported that some states could use more assistance in creating valid academic assessments for students with limited English proficiency. While 4 of the 5 states we studied in depth had significant experience in, and multiple staff devoted to, developing language arts and mathematics assessments, some members of our expert group pointed out that the assessment departments in other states have limited resources and expertise, as well as high turnover. As a result, these states need help to conduct appropriate analyses that will offer useful information about the validity and reliability of their academic assessments for students with limited English proficiency. An Education official told us that the agency recently began offering technical assistance to states that need help addressing issues raised during their peer reviews.

Our group of experts suggested several areas where states could benefit from additional assistance and guidance in developing academic assessments for students with limited English proficiency. Several members noted the lack of good research on what kinds of accommodations can help mitigate language barriers for students with limited English proficiency. Several experts also believed that some states need more information on how to implement universal design principles to develop assessments that produce valid results for students with limited English proficiency. In addition, some group members pointed out that developing equivalent assessments in other languages (that is, assessments that measure the same thing and are of equivalent difficulty) is challenging and that states need more information about how to develop such assessments, as well as examples.

Education Has Offered Different Accountability Options for Students with Limited English Proficiency, but State Officials Reported Additional Flexibility Is Needed

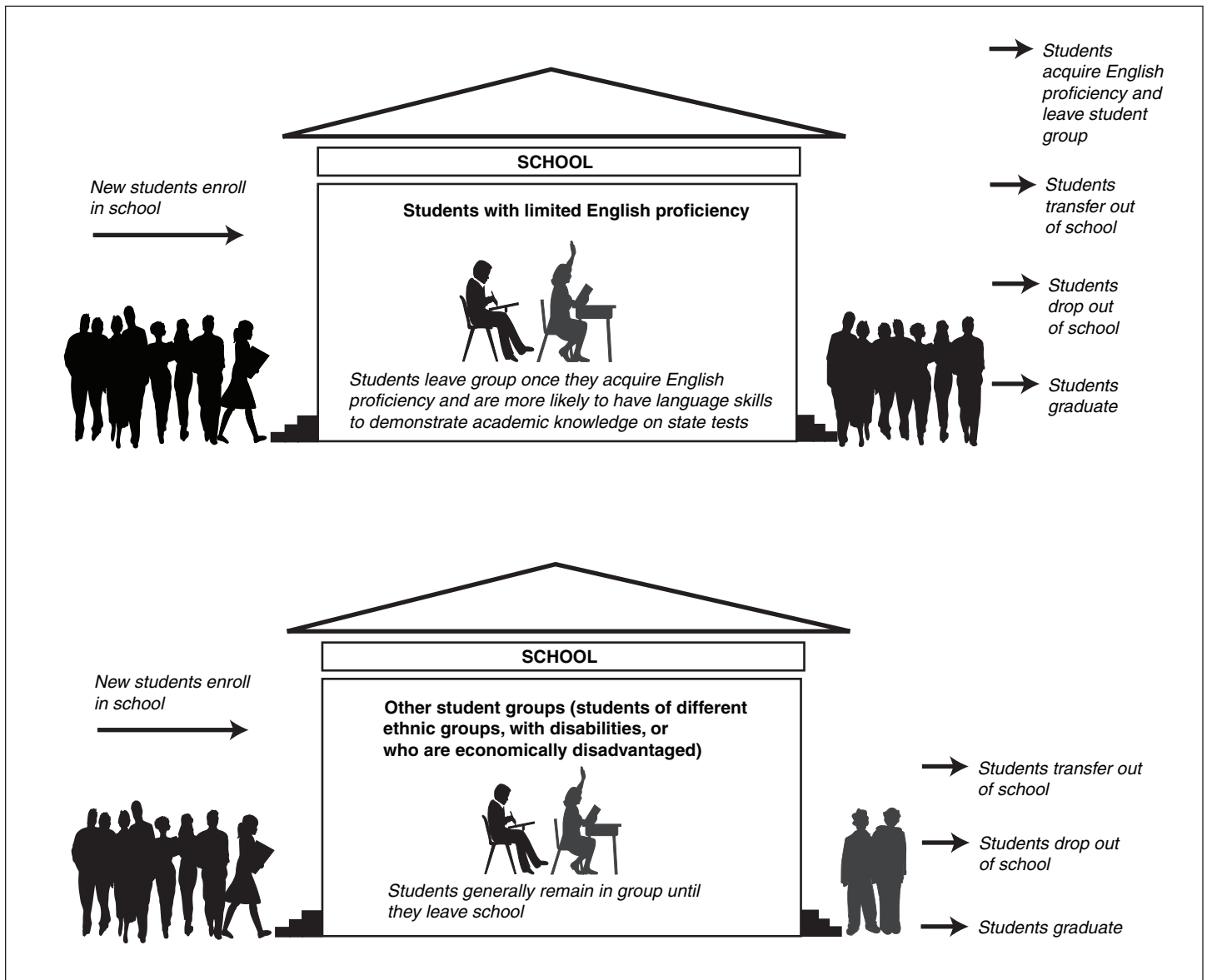
Education has offered states several flexibilities in tracking academic progress goals for students with limited English proficiency to support their efforts to develop appropriate accountability systems for these students. In a February 2004 notice, Education recognized the existence of language barriers that hinder the assessment of students who have been in the country for a short time and provided some testing flexibility for these students. Specifically, Education does not require students with limited English proficiency to participate in a state's language arts assessment during their first year in U.S. schools. In addition, while these students must take a state's mathematics assessment during their first year in U.S.

schools, a state may exclude their scores in determining whether it met its progress goals.²⁶

Education offered additional flexibility in its February 2004 notice, recognizing that limited English proficiency is a more transient quality than having a disability or being of a particular race. Unlike the other NCLBA student groups, students who achieve English proficiency leave the group at the point when they are more prepared to demonstrate their academic knowledge in English, while new students with lower English proficiency are constantly entering the group (see fig. 6). Given the group's continually changing composition, meeting progress goals may be more difficult than doing so for other student groups, especially in districts serving large numbers of students with limited English proficiency. To compensate for this, Education allowed states to include, for up to 2 years, the scores of students who were formerly classified as limited English proficient when determining whether a state met its progress goals for students with limited English proficiency. In addition, Education has approved requests from several states to permit students who have been redesignated as English proficient to remain in the group of students with limited English proficiency until they have achieved the proficient level on the state's language arts assessment for 1 or more years.

²⁶On June 24, 2004, Education issued proposed regulations on these flexibilities for students with limited English proficiency for public comment, but the regulations have not been finalized as of June 2006.

Figure 6: Movement of Students In and Out of Limited English Proficient Student Group and Other Student Groups



Source: GAO analysis and Art Explosion images.

Several state and local officials in our study states told us that additional flexibility would be helpful to ensure that the annual progress measures provide meaningful information about the performance of students with limited English proficiency. Officials in 4 of the states we studied suggested that certain students with limited English proficiency should be

exempt for longer periods from taking academic content assessments or that their test results should be excluded from a state's annual progress determination for a longer period than is currently allowed. Several officials voiced concern that some of these students have such poor English skills or so little previous school experience that the assessment results do not provide any meaningful information. Instead, some of these officials stated that students with limited English proficiency should not be included in academic assessments until they demonstrate appropriate English skills on the state's English language proficiency assessment. However, the National Council of La Raza, an Hispanic advocacy organization, has voiced concern that excluding too many students with limited English proficiency from a state's annual progress measures will allow some states and districts to overlook the needs of these students. Education officials reported that they are developing a regulation with regard to how test scores for this student group are included in a state's annual progress measures, but it has not yet been finalized.

With respect to including the scores of students previously classified as limited English proficient in a state's progress measures for this group for up to 2 years, officials in 2 of our 5 study states, as well as one member of our expert group, thought it would be more appropriate for these students to be counted in the limited English proficient group throughout their school careers—but only for accountability purposes. They pointed out that by keeping students formerly classified as limited English proficient in the group, districts that work well with these students would see increases in the percentage who score at the proficient level in language arts and mathematics. An Education official explained that the agency does not want to label these students as limited English proficient any longer than necessary and considered including test results for these students for 2 years after they have achieved English proficiency to be the right balance. Education officials also noted that including all students who were formerly limited English proficient would inflate the achievement measures for the student group.

District officials in 4 of the states we studied argued that tracking the progress of individual students in this group is a better measure of how well these students are progressing academically. Officials in one district pointed to a high school with a large percentage of students with limited English proficiency that had made tremendous progress with these students, doubling the percentage of students achieving academic proficiency. The school missed the annual progress target for this group by a few percentage points, but school officials said that the school would be considered successful if it was measured by how much individual

students had improved in their test scores. A district official in another state explained that many students with limited English proficiency initially have very low test scores, but demonstrate tremendous improvement in these scores over time. In response to educators and policymakers who believe such an approach should be used for all students, Education initiated a pilot project in November 2005, allowing a limited number of states to incorporate measures of student progress over time in determining whether districts and schools met their annual progress goals. Even using this approach, however, states must still establish annual goals that lead to all students achieving proficient scores by 2014.²⁷

Conclusions

NCLBA has focused attention on the academic performance of all students, especially those who have historically not performed as well as the general student population, such as students with limited English proficiency. NCLBA requires states to include these students in their language arts and mathematics assessments and to assess them in a valid and reliable manner, and states are in various stages of doing so. Although Education has provided some technical assistance to states, our group of experts and others have noted the complexity of developing academic assessments for these students and have raised concerns about the technical expertise of states to ensure the validity and reliability of assessment results. Using assessment results that are not a good measure of student knowledge is likely to lead to poor measures of state and district progress, thereby undermining NCLBA's purpose to hold schools accountable for student progress. Further, although most states offered these students accommodations, research on their appropriateness is limited. National research on accommodations has informed states' practices in assessing students with disabilities. Without similar research efforts, accommodations offered to students with limited English proficiency may not improve the validity of their test results.

While Education has provided some support and training to states, officials in a number of states are still uncertain about how to comply with some of the more technical requirements of the new English language proficiency assessments required by NCLBA. State officials reported that

²⁷See GAO, *No Child Left Behind Act: States Face Challenges Measuring Academic Growth That Education's Initiatives May Help Address*, [GAO-06-661](#) (Washington, D.C.: July 17, 2006) for further information on Education's pilot project.

they need more guidance from Education to develop these assessments. States have had to develop many new assessments under NCLBA for both English language proficiency and academic content, and some states may lack the technical expertise to develop assessments that produce valid results for students with limited English proficiency. Without more specific guidance outlining Education's requirements, states may spend time developing English language proficiency assessments that do not adequately track student progress in learning English or otherwise meet NCLBA's requirements.

Including students with limited English proficiency in NCLBA's accountability framework presents unique challenges. For example, students who have little formal schooling may make significant progress in learning academic skills, but may not achieve proficiency on state academic assessments for several years. The movement of students into and out of the group also makes it more difficult for the group to meet state progress goals, even when these students are making academic progress. Education has addressed some of the unique characteristics of this student group and provided some flexibility in how states and districts are held accountable for the progress of these students. However, these current flexibilities may not fully account for the characteristics of certain students with limited English proficiency, such as those who have little previous formal schooling.

Recommendations for Executive Action

We recommend that the Secretary of Education

1. Support additional research on appropriate accommodations for students with limited English proficiency and disseminate information on research-based accommodations to states.
2. Determine what additional technical assistance states need with respect to assessing the academic knowledge of students with limited English proficiency and to improve the validity and reliability of their assessment results (such as consultations with assessment experts and examples of assessments targeted to these students) and provide such additional assistance.
3. Publish additional guidance with more specific information on the requirements for assessing English language proficiency and tracking the progress of students with limited English proficiency in learning English.

-
4. Explore ways to provide additional flexibilities to states in terms of holding states accountable for students with limited English proficiency. For example, among the flexibilities that could be considered are
 - allowing states to include the assessment scores for all students formerly considered to have limited English proficiency in a state's annual progress results for the group of students with limited English proficiency,
 - extending the period during which the assessment scores for some or all students with limited English proficiency would not be included in a state's annual progress results, and
 - adjusting how states account for recent immigrants with little formal schooling in their annual progress results.

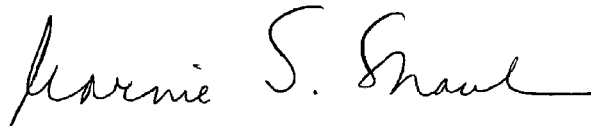
Agency Comments

We provided a draft of this report to Education for review and comment. The agency provided comments, which are reproduced in appendix VII. Education also provided technical clarifications, which we incorporated when appropriate. Education agreed with our first three recommendations. The department noted that it has conducted some research on the effectiveness of accommodations and is currently working with its National Research and Development Center for Assessment and Accountability to synthesize the existing research literature on the assessment of students with limited English proficiency. Education also explained that it has begun the process of identifying the additional technical assistance needs of states with respect to academic assessments; specifically, it will have its Assessment and Accountability Comprehensive Center conduct a needs assessment this fall to determine specific areas in which states need assistance and will provide technical assistance to address those areas. In addition, the department stated that it is exploring ways to help states assess English language proficiency.

Education did not explicitly agree or disagree with our fourth recommendation. Instead, the agency commented that it has explored and already provided various types of flexibility regarding the inclusion of students with limited English proficiency in accountability systems. Further, Education noted that it is in the process of completing a regulation on flexibility for these students. However, the department also emphasized that all students with limited English proficiency must be included in school accountability systems to improve both instruction and achievement outcomes. Through our recommendation, we encourage the department to continue its efforts.

We are sending copies of this report to the Secretary of Education, relevant congressional committees, and other interested parties. We will make copies available to others upon request. In addition, the report will be available at no charge on GAO's Web site at <http://www.gao.gov>.

If you or your staff have any questions or wish to discuss this report further, please contact me at (202) 512-7215 or at shaulm@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. Other contacts and major contributors are listed in appendix VIII.

A handwritten signature in black ink that reads "Marnie S. Shaul". The signature is written in a cursive style with a long horizontal flourish at the end.

Marnie S. Shaul
Director, Education, Workforce,
and Income Security Issues

Appendix I: GAO's Group of Experts on Assessing the Academic Knowledge of Students with Limited English Proficiency

On January 20, 2006, GAO, with the assistance of the National Academy of Sciences, convened a group of experts in Davis, California, to discuss issues related to assessing the academic knowledge of students with limited English proficiency. Specifically, we asked the group to discuss the following questions:

- To meet the requirements of the No Child Left Behind Act (NCLBA), what steps should states take to ensure the validity and reliability of language arts and mathematics assessments for students with limited English proficiency?
- What steps should states take to ensure that students with limited English proficiency receive appropriate accommodations on language arts and mathematics assessments?
- Given NCLBA's accountability framework, what is the most appropriate way to hold schools and districts accountable for the performance of students with limited English proficiency?
- How can the U.S. Department of Education assist states in their efforts to meet NCLBA's assessment and accountability requirements for students with limited English proficiency?

Group members who were selected had significant technical and research expertise in assessments issues. Some members had technical expertise on general assessment issues, while others had specifically conducted assessment research focused on students with limited English proficiency. The members of our expert group are listed below:

Dr. Jamal Abedi
University of California, Davis

Dr. Stephen Dunbar
The University of Iowa

Dr. Richard Durán
University of California, Santa Barbara

Dr. Steven Ferrara
American Institutes for Research

Dr. Patricia Gándara
University of California, Davis

**Appendix I: GAO's Group of Experts on
Assessing the Academic Knowledge of
Students with Limited English Proficiency**

Dr. Edward Haertel
Stanford University

Dr. Rebecca Kopriva
University of Maryland

Dr. Stanley Rabinowitz
WestEd

Dr. Charlene Rivera
The George Washington University

Dr. Rebecca Zwick
University of California, Santa Barbara

Appendix II: Determining Adequate Yearly Progress for Student Groups

NCLBA requires states to report adequate yearly progress (AYP) results at the state level for each of the required student groups, including students with limited English proficiency. The law also requires Education, starting in the 2004-2005 school year, to make an annual determination about whether states have made adequate yearly progress for each student group.¹ Education has issued some general regulations regarding state-level adequate yearly progress. However, Education has not yet collected any such state-level adequate yearly progress results and has not issued any guidance on how states should determine whether a student group has made adequate yearly progress. As a result, some states have not yet made adequate yearly progress determinations for student groups at the state level.

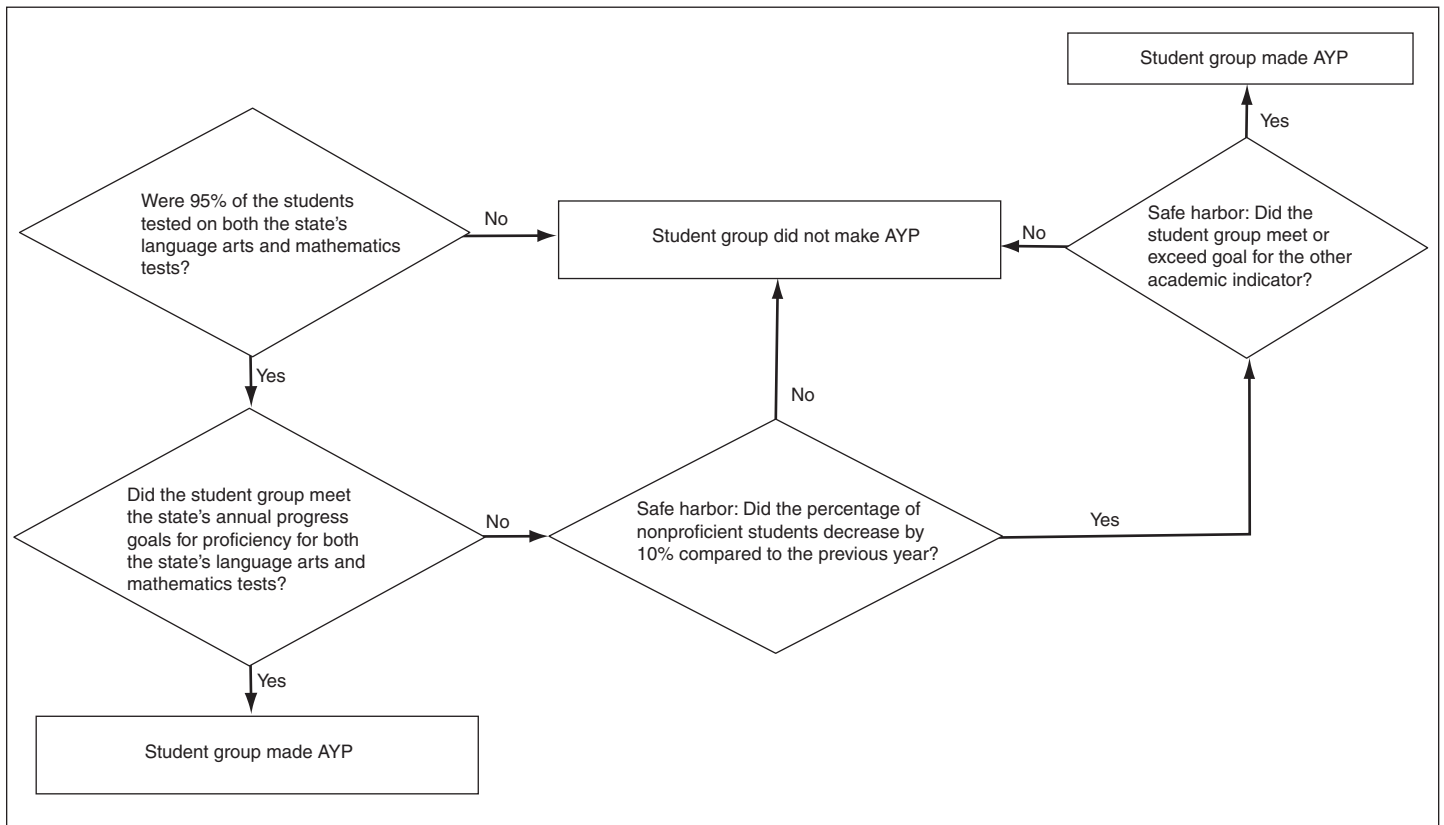
In order for a student group, such as students with limited English proficiency, to make adequate yearly progress, it must make a number of different goals. Specifically:

- At least 95 percent of students in the group must take the state's language arts and mathematics assessments, and
- The student group must meet the progress goals established by the state for both language arts and mathematics proficiency or
- The percentage of students who did not achieve proficient scores must have decreased by at least 10 percent from the previous year, and the student group must also meet the progress goals established by the state for its other academic indicator (graduation rate for high schools and usually attendance rate for other schools).

Figure 7 illustrates the basic decision process for determining adequate yearly progress for a student group.

¹NCLBA also requires Education to annually determine whether states have met their Title III goals related to increases in students making progress in learning English and attaining English proficiency.

Figure 7: Process for Determining Adequate Yearly Progress for a Student Group

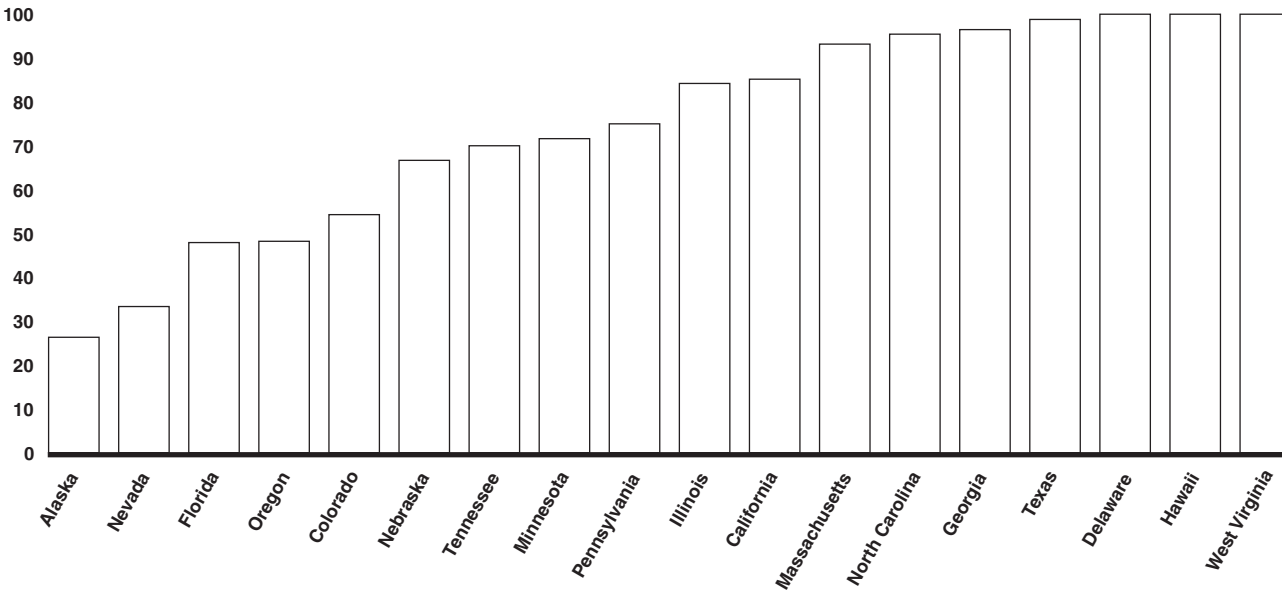


Source: GAO analysis.

Because states have different assessment systems, they use different methods for determining adequate yearly progress. A state can have an assessment system that allows it to create the same progress goal for mathematics and language arts for all grades, despite using different tests in each grade. In this case, the state could review data for all students in a student group across the state to determine if the group met its annual progress goals. A state can also establish different progress goals for different grades or groups of grades, depending on the particular test being used. In this case, according to an Education official, a state would have to meet all the proficiency and participation goals for all the different grades or groups of grades in order to make adequate yearly progress.

Appendix III: Percentage of Districts Making AYP Goals for Mathematics for Students with Limited English Proficiency

Percentage of districts meeting mathematics progress goals



Source: GAO analysis of district report cards and district data provided by state officials.

Notes: Data are for school year 2003-2004.

We requested district-level achievement data from 20 states, and 18 states responded to our request.

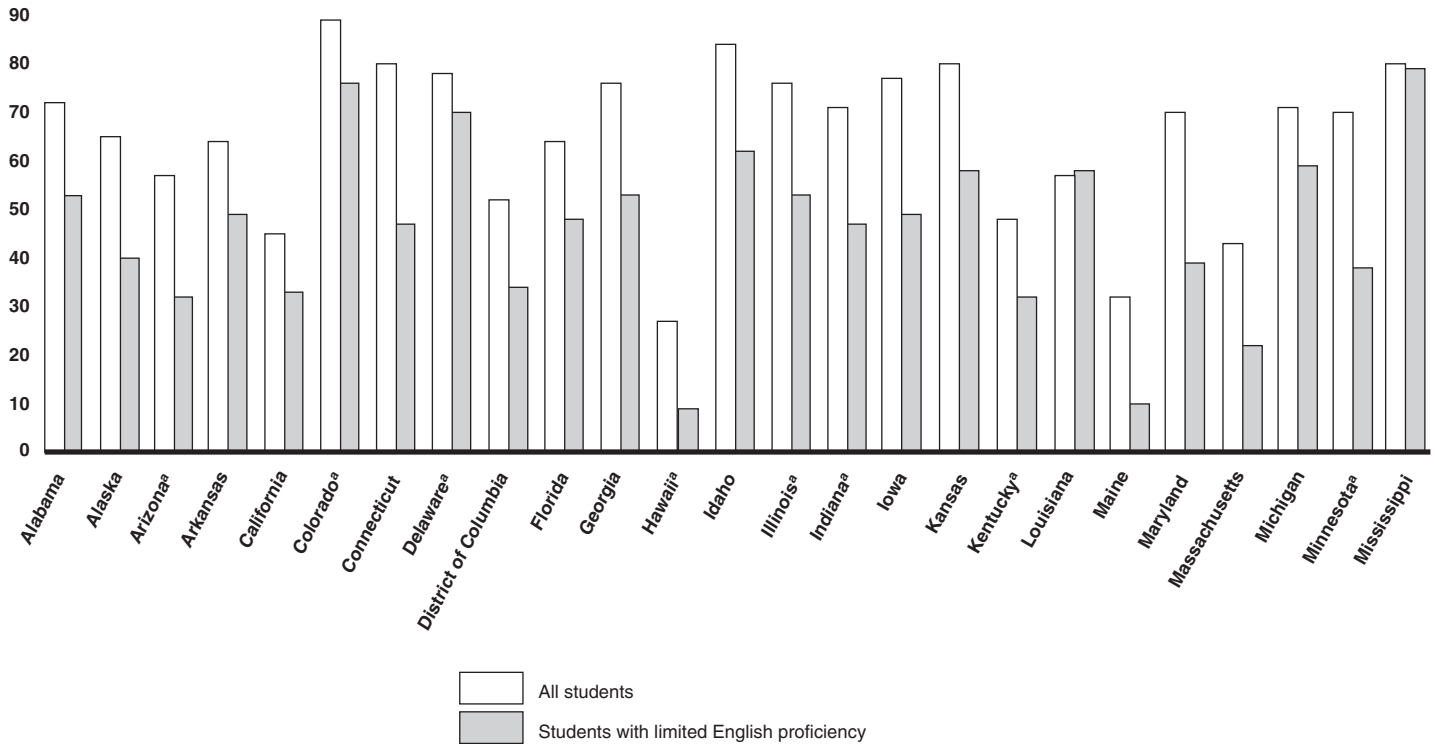
When districts reported proficiency data for different grades or groups of grades, we determined that the percentage of students with limited English proficiency met a state's mathematics progress goal if the student group met the goal for all grades reported.

Results from charter schools are included when a charter school is its own school district or part of a larger school district.

Hawaii only has one school district.

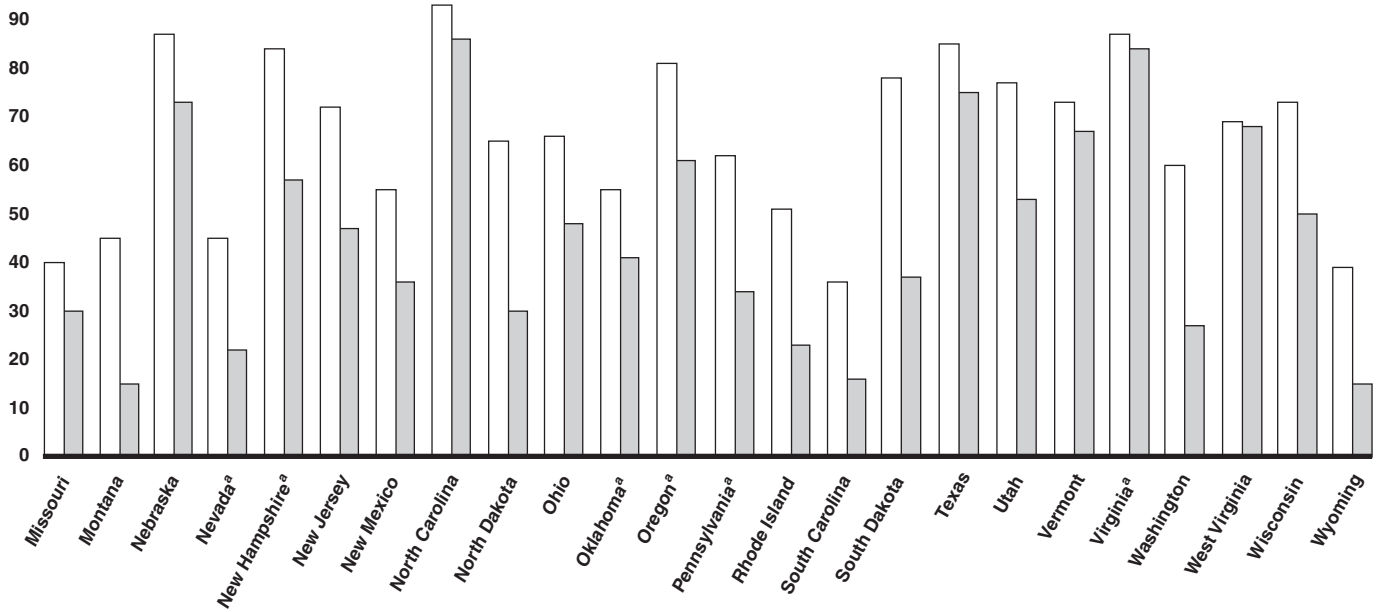
Appendix IV: Proficiency Scores on Mathematics Tests for All Students and Students with Limited English Proficiency

Percentage of elementary students achieving proficient scores



Source: Consolidated performance reports for the 2003-2004 school year.

Appendix IV: Proficiency Scores on Mathematics Tests for All Students and Students with Limited English Proficiency



Notes: New York and Tennessee did not provide assessment data in their state consolidated performance reports for the 2003-2004 school year.

The results for Arkansas do not include those students with limited English proficiency who were considered proficient based on the state's portfolio assessment.

The total student population includes students with limited English proficiency.

^aMost states reported assessment data for students in the fourth grade. States marked with this superscript reported on some other grade at the elementary school level, usually either third grade or fifth grade.

Appendix V: Enhanced Assessment Consortia Participation

	World-Class Instructional Design and Assessment (WIDA) Consortium	State Collaborative on Assessment and Student Standards (SCASS) Consortium	Mountain West Assessment Consortium (MWAC)	Pennsylvania Enhanced Assessment Grant (PA EAG)
Assessment	Assessing Comprehension and Communication in English State-to-State for English Language Learners (ACCESS for ELLs)	English Language Development Assessment (ELDA)	MWAC	Comprehensive English Language Learning Assessment (CELLA)
Consortia states				
Using assessment in 2005-2006 school year	Alabama Delaware District of Columbia Georgia Illinois Maine New Hampshire New Jersey Oklahoma Rhode Island Vermont Wisconsin	Iowa Louisiana Nebraska Ohio South Carolina West Virginia	Idaho ^a New Mexico ^a Michigan ^{a b}	Tennessee
Not using assessment in 2005-2006 school year	none	Kentucky Nevada North Carolina Texas	Alaska Colorado Montana Nevada North Dakota Utah Wyoming	Florida Maryland Michigan ^b Pennsylvania

Source: Interviews with consortia and state officials.

Note: This table reflects states that participated in the consortia prior to or during the 2005-2006 school year. Some states are no longer consortia members.

^aUsing test items from consortium-developed assessment.

^bParticipated in more than one consortium.

Appendix VI: English Language Proficiency Assessments Used in the 2005-2006 School Year, by State

State	English Language Proficiency Assessment
Alabama	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
Alaska	IDEA Proficiency Test
Arizona	Stanford English Language Proficiency Test
Arkansas	MAC II (Maculaitis Assessment of Competencies) Test of English Language Proficiency
California	California English Language Development Test
Colorado	Colorado English Language Assessment
Connecticut	LAS (Language Assessment System) Links
Delaware	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
District of Columbia	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
Florida	Various off-the-shelf tests ^a
Georgia	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
Hawaii	LAS (Language Assessment System) Links
Idaho	Mountain West Assessment Consortium test items
Illinois	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
Indiana	LAS (Language Assessment System) Links
Iowa	English Language Development Assessment (SCASS)
Kansas	Kansas English Language Proficiency Assessment
Kentucky	2004 IDEA Proficiency Test or Language Assessment Scales (LAS) ^a
Louisiana	English Language Development Assessment (SCASS)
Maine	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
Maryland	LAS (Language Assessment System) Links
Massachusetts	Massachusetts English Proficiency Assessment
Michigan	English Language Proficiency Assessment ^b (includes Mountain West Consortium test items)
Minnesota	Test of Emerging Academic English, Minnesota Student Oral Language Observation Matrix, and checklist for reading and writing for K-2 students
Mississippi	Stanford English Language Proficiency Test
Missouri	MAC II (Maculaitis Assessment of Competencies) Test of English Language Proficiency
Montana	Iowa Test of Basic Skills, Woodcock-Muñoz Language Survey (English), or other state-approved test ^a
Nebraska	English Language Development Assessment (SCASS)

Appendix VI: English Language Proficiency Assessments Used in the 2005-2006 School Year, by State

State	English Language Proficiency Assessment
Nevada	LAS (Language Assessment System) Links
New Hampshire	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
New Jersey	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
New Mexico	New Mexico English Language Proficiency Assessment (includes Mountain West Consortium test items)
New York	New York State English as a Second Language Achievement Test
North Carolina	IDEA Proficiency Test
North Dakota	2004 IDEA Proficiency Test, Woodcock-Muñoz Language Survey (English), and Language Assessment Scales (LAS) ^a
Ohio	English Language Development Assessment (SCASS)
Oklahoma	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
Oregon	English Language Proficiency Assessment ^b
Pennsylvania	Stanford English Language Proficiency Test
Rhode Island	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
South Carolina	English Language Development Assessment (SCASS)
South Dakota	Dakota English Language Proficiency assessment
Tennessee	Comprehensive English Language Learning Assessment (PA EAG)
Texas	Texas English Language Proficiency Assessment System; consists of Reading Proficiency Tests in English and Texas Observation Protocols
Utah	2004 IDEA Proficiency Test
Vermont	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
Virginia	Stanford English Language Proficiency Test
Washington	Washington Language Proficiency Test
West Virginia	English Language Development Assessment (SCASS)
Wisconsin	Assessing Comprehension and Communication in English State-to-State for English Language Learners (WIDA)
Wyoming	Wyoming English Language Learner Assessment

Source: Interviews with consortia and state officials.

^aState allows school districts to individually choose tests.

^bAssessments are not the same; Oregon's is a state developed assessment, while Michigan's is a combination of an augmented version of the Stanford English Language Proficiency Test and test items from the Mountain West Assessment Consortium.

Appendix VII: Comments from the Department of Education



UNITED STATES DEPARTMENT OF EDUCATION

OFFICE OF ELEMENTARY AND SECONDARY EDUCATION

July 20, 2006

THE ASSISTANT SECRETARY

Ms. Marnie S. Shaul
Director, Education, Workforce
and Income Security Issues
Government Accountability Office
441 G Street, NW
Washington, DC 20548

Dear Ms. Shaul:

I am writing in response to your request for comments on the Government Accountability Office (GAO) draft report (GAO-06-815), dated August 2006, and entitled "No Child Left Behind Act: Assistance from Education Could Help States Better Measure Progress of Students with Limited English Proficiency." I appreciate the opportunity to comment on the draft report and provide insight on our activities to support states in serving limited English proficient (LEP) students.

The following are responses to the recommendations in the report to assist states better measure the progress of students with limited English proficiency:

Recommendation 1. Support additional research on appropriate accommodations for students with limited English proficiency and disseminate information on research-based accommodations to states.

The Department agrees with acquiring additional information on better serving LEP students. The Department's Institute of Education Sciences (IES), through its National Assessment of Educational Progress (NAEP), has sponsored studies of the effectiveness of various accommodations for LEP students, such as simplified English and use of translations, glossaries, and bilingual dictionaries, and we have applied the findings to NAEP. The Department has also shared the results of the studies at various meetings and national conferences. NAEP has also conducted studies on the impact on validity of test results of providing accommodations to LEP students, and on the impact of excluding such students from the assessment.

In addition, IES is working with its National Research and Development Center for Assessment and Accountability -- known as the Center for Research on Evaluation, Standards, and Student Testing (CRESST) -- to synthesize the existing research literature on the assessment of LEP students. The synthesis addresses both accessibility and validity. After disseminating the information to states, CRESST will work with states to help them review and improve their assessment instruments and procedures, especially with regard to LEP students.

400 MARYLAND AVE., S.W., WASHINGTON, DC 20202
www.ed.gov

Our mission is to ensure equal access to education and to promote educational excellence throughout the nation.

Recommendation 2. Determine what additional technical assistance states need with respect to assessing the academic knowledge of students with limited English proficiency and to improve the validity and reliability of the assessment results (such as consultations with assessment experts and examples of assessments targeted to these students) and provide such additional assistance.

The Department has begun the process of identifying the additional technical assistance needs of states related to assessing the academic knowledge of limited English proficient students. For this process, the Department is using the resources of our Assessment and Accountability Comprehensive Center (AACC) and the 16 Regional Comprehensive Centers (RCCs) funded by the Department. Specifically, the AACC has developed a plan approved by the Department to provide information and resources to the RCCs, and through the RCCs to State education agencies (SEAs), regarding the assessment and accountability of their special student populations (LEP students and students with disabilities). During 2006-07 and subsequent years, the AACC will continue to address aspects of assessment and accountability as they relate to LEP students. This fall, the AACC will conduct a needs assessment of the RCCs and SEAs to determine specific areas in which SEAs need assistance and, subsequently, the AACC will use that information to develop resource materials and to provide specific technical assistance. We anticipate that the AACC, in coordination with the Department, will be able to provide the following resources to the RCCs and SEAs during the 2006-07 school year:

1. A summary evaluation of evidence related to the technical aspects of content or language proficiency assessments for LEP students (including the reliability and validity of widely used instruments), and a synopsis of information on existing available resources and on those being developed. The evaluation will yield results that are intended to assist consumers of assessments (e.g., SEAs, LEAs) for LEP students. It will focus on the technical adequacy (i.e., validity, reliability, freedom from bias) of evidence related to assessments used to meet relevant Title I and Title III requirements under the Elementary and Secondary Education Act, as amended by the No Child Left Behind Act (ESEA) (e.g., assessments developed specifically for a state, consortium-developed assessments, assessments published for wider use). Technical evidence associated with assessments for LEP students will be evaluated against a comprehensive set of validated criteria (Rabinowitz & Sato, 2005). Using the results of this evaluation, RCCs and SEAs should be better able to gauge the technical adequacy of the assessments they are using or have available for use with their LEP students and to identify appropriate and necessary next steps for ensuring the assessments' validity and, ultimately, the defensibility of their assessment systems and results.
2. Guidelines for the interpretation of regulations regarding the implementation of assessments for LEP students (under Titles I and III of ESEA), including other pertinent statutes and regulations. These guidelines are expected to address compliance with NCLB regulations for special student populations such as LEP students and will provide information to RCCs as they work with their states in gauging whether they are meeting federal requirements; will focus attention on priority issues related to implementing practices and systems that are in

compliance with federal regulations; and will select specific implementation strategies, given the particular needs and conditions of the state. Selected strategies will have evidence of effectiveness in a state or states and will be examples of methods for enhancing and supporting federal guidance.

3. Web-posting and dissemination of evidence-based products and services for easy access to the most current information.

The AACC, headed by a national expert on assessment technical quality and staffed by a team of academicians and former practitioners, is well suited to conduct the needs assessment necessary to develop resources and provide technical assistance. The Department will monitor the development of the needs assessment and the subsequent activities of the AACC to ensure that the information being developed is of consistently high quality and relevance.

Recommendation 3. Publish additional guidance with more specific information on the requirements for assessing English language proficiency and tracking the progress of students with limited English proficiency in learning English.

The Department agrees with this recommendation and is exploring ways -- through guidance to states, the Office of English Language Acquisition's annual Title III Summit, and other means -- to help states appropriately assess and track the progress of students with limited proficiency in English.

Recommendation 4. Explore ways to provide additional flexibilities to states in terms of holding states accountable for students with limited English proficiency. Such flexibilities could include, for example.

- Allowing states to include the assessment scores for all students formerly considered to have limited English proficiency in a state's annual progress results for the group of students with limited English proficiency,
- Extending the period during which the assessment scores for some or all students with limited English proficiency would not be included in a state's annual progress results, and
- Adjusting how states account for recent immigrants with little formal schooling in their annual progress results.

As your report points out, the Department has explored and has already provided various types of flexibility to states regarding LEP students and how they are considered in states' annual progress results. The Department has engaged in discussions with researchers, practitioners, and educators to explore appropriate means of providing appropriate flexibility in holding schools accountable for the academic progress of children who have not grown up speaking English.

NCLB gives states some flexibility in defining the students who constitute the LEP subgroup. States also have some flexibility in determining how to assess their LEP

students. States can offer a menu of accommodations (e.g., use of bilingual dictionaries, extra time, use of translators) or develop alternative assessments (e.g., a native-language version or simplified English version of its assessment). The law also allows states three years to test LEP students in language arts using a native language assessment, with an additional two years if needed on a case-by-case basis.

The Department recognizes that LEP students new to the United States often have some challenges in participating in state assessments due to language barriers and the challenge of adjusting to U.S. schools. In addition, because students exit the LEP subgroup once they attain English language proficiency, states may have difficulty demonstrating improvements on state assessments for these students.

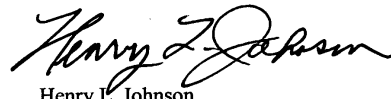
To address these two issues, the Department has received public comments on a notice of proposed rulemaking and is in the process of completing a regulation on flexibility for “recent arrivals” and “formerly LEP” students. In the meantime, the Department is implementing a transitional policy with regard to these policy issues. Under that transitional policy, recently arrived LEP students during the first year of enrollment in U.S. schools can be exempted from taking reading/language arts assessments and the scores of recently arrived LEP students on state math assessments, as well as reading/language arts assessments if taken, can be excluded from adequate yearly progress (AYP) calculations for that first year in U.S. schools as well. In addition, the flexibility allows states to include, within the LEP subgroup for AYP purposes, for up to two years, the scores on state assessments for “formerly LEP” students who have since attained English proficiency.

While the Department has explored and implemented flexibility policies related to LEP students, we also know that regular and high-quality student assessment, school accountability, and clear and accessible data and information about student and school performance are absolutely essential to improving instruction for LEP students. The No Child Left Behind Act shines a bright light on the needs of students who have so often been left behind in our Nation’s schools – including students with limited English proficiency. Therefore, the Department believes that all LEP students need to be included and visible in our school accountability systems to improve both instruction and achievement outcomes. The ability to better measure performance and analyze student data is an important vehicle for improving instruction and closing the achievement gap for LEP students and reaching the goal of the No Child Left Behind Act of 100 percent of students being able to master grade level work by 2013-14.

**Appendix VII: Comments from the
Department of Education**

We appreciate the opportunity to share our comments, accomplishments and plans. Please let me know if you need additional information regarding activities underway at the Department to assist States better measure the progress of students with limited English proficiency.

Sincerely,

A handwritten signature in cursive script that reads "Henry L. Johnson".

Henry L. Johnson

Enclosure

Appendix VIII: GAO Contacts and Acknowledgments

GAO Contacts

Marnie Shaul, (202) 512-7215, shaulm@gao.gov and Cornelia Ashby (202) 512-7215, ashbyc@gao.gov

Staff Acknowledgments

Harriet Ganson (Assistant Director) and Michelle St. Pierre (Analyst-in-Charge) managed all aspects of this assignment. Shannon Groff, Eileen Harrity, and Krista Loose made significant contributions to this report. Katie Brillantes contributed to the initial design of the assignment. Carolyn Boyce, John Mingus, and Lynn Musser provided key technical support, James Rebbe provided legal support, and Scott Heacock assisted in message and report development.

Related GAO Products

No Child Left Behind Act: States Face Challenges Measuring Academic Growth That Education's Initiatives May Help Address. [GAO-06-661](#). Washington, D.C.: July 17, 2006.

No Child Left Behind Act: Improved Accessibility to Education's Information Could Help States Further Implement Teacher Qualification Requirements. [GAO-06-25](#). Washington, D.C.: November 21, 2005.

No Child Left Behind Act: Education Could Do More to Help States Better Define Graduation Rates and Improve Knowledge about Intervention Strategies. [GAO-05-879](#). Washington, D.C.: September 20, 2005.

No Child Left Behind Act: Most Students with Disabilities Participated in Statewide Assessments, but Inclusion Options Could Be Improved. [GAO-05-618](#). Washington, D.C.: July 20, 2005.

Head Start: Further Development Could Allow Results of New Test to Be Used for Decision Making. [GAO-05-343](#). Washington, D.C.: May 17, 2005.

No Child Left Behind Act: Education Needs to Provide Additional Technical Assistance and Conduct Implementation Studies for School Choice Provision. [GAO-05-7](#). Washington, D.C.: December 10, 2004.

No Child Left Behind Act: Improvements Needed in Education's Process for Tracking States' Implementation of Key Provisions. [GAO-04-734](#). Washington, D.C.: September 30, 2004.

GAO's Mission

The Government Accountability Office, the audit, evaluation and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site (www.gao.gov). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to www.gao.gov and select "Subscribe to Updates."

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are \$2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. Government Accountability Office
441 G Street NW, Room LM
Washington, D.C. 20548

To order by Phone: Voice: (202) 512-6000
TDD: (202) 512-2537
Fax: (202) 512-6061

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Gloria Jarmon, Managing Director, JarmonG@gao.gov (202) 512-4400
U.S. Government Accountability Office, 441 G Street NW, Room 7125
Washington, D.C. 20548

Public Affairs

Paul Anderson, Managing Director, AndersonP1@gao.gov (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, D.C. 20548