CRESST POLICY BRIEF 7
Winter 2004

Challenges in the No Child Left Behind Act for English Language Learners

Jamal Abedi and Ron Dietel[1]
National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/UCLA

Jamal Abedi

Jamal Abedi is a CRESST senior researcher and a faculty member at the UCLA Graduate School of Education & Information Studies. His recent studies have focused on the impact of linguistic factors and accommodations for English language learners.

Ron Dietel

Ron Dietel is the CRESST assistant director for research use and communications.

The No Child Left Behind Act of 2001 (NCLB) requires that all children, including English language learners (ELLs), reach high standards by demonstrating proficiency in English language arts and mathematics by 2014. Schools and districts must help ELL students, among other subgroups, make continuous progress toward this goal, as measured by performance on state tests, or risk serious consequences.

Through these mandates, NCLB establishes high expectations for all students and seeks to reduce the achievement gap between advantaged and disadvantaged students. These are worthy goals, which require extraordinary improvement in student learning. The challenges for English language learners are especially difficult, involving both educational and technical issues, including:

1. **Historically low ELL performance and slow improvement.** State tests show that ELL students' school performance is far below that of other students, oftentimes 20 to 30 percentage points, and usually shows little improvement across many years.

2. **Measurement accuracy.** CRESST research shows that the language demands of tests negatively influence accurate measurement of ELL performance. For the ELL student, tests measure both achievement *and* language ability.

3. **Instability of the ELL student subgroup.** The goal of redesignating high-performing ELL students as language-proficient students causes ELL high achievers to exit the ELL subgroup. The consequence is downward pressure on ELL test scores worsened by the addition of new ELL students, who are typically low achieving.

4. **Factors outside of a school's control.** CRESST research shows substantial nonschool effects on student learning even within ELL subgroups. Schools are therefore unable to control all factors related to student achievement.

These ELL issues are described in the remainder of this brief, together with some suggestions to help schools meet the NCLB goals. Our comments are based on a series of research reports by Jamal Abedi and others, listed in the References and Resources section.

Historically Low ELL Performance and Slow Improvement

CRESST research, supported by the National Assessment of Educational Progress (NAEP) and state test results, shows that English language learners consistently perform lower than other students and frequently lower than many other subgroups. The

continued on next page

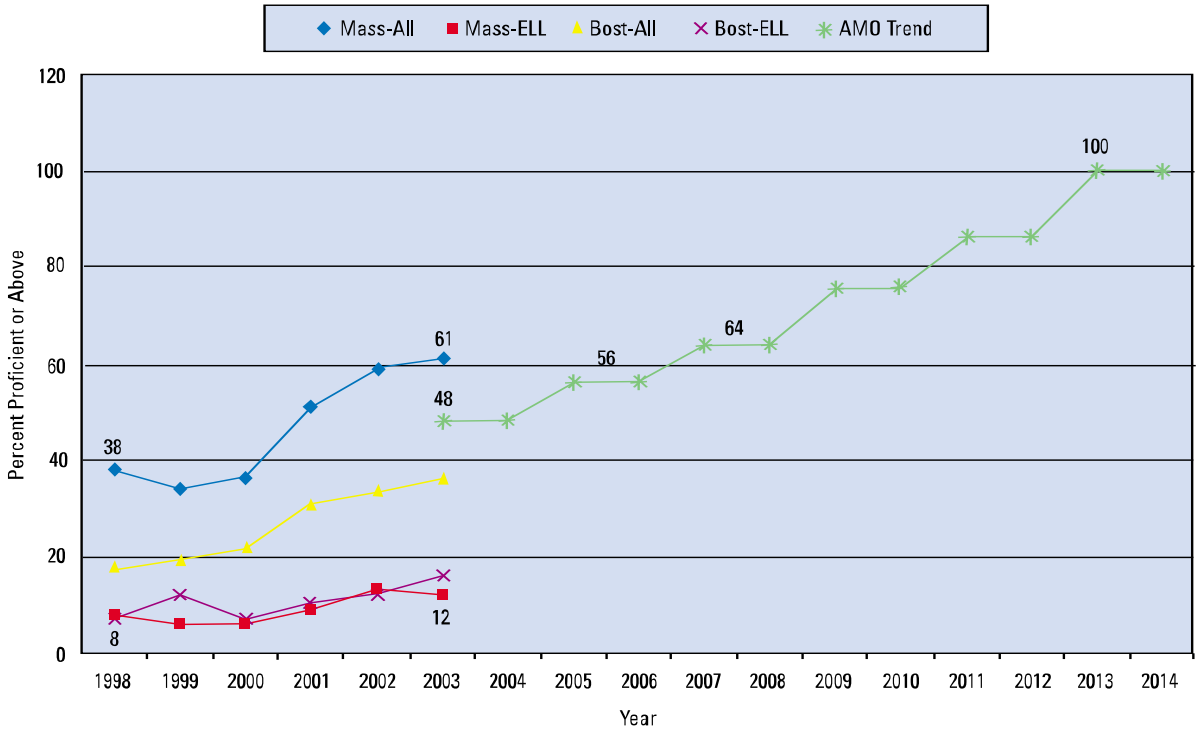

Figure 1. MCAS 10th-grade English language arts assessment 1998-2003 scores and annual measurable objectives (AMO) trend line.

Massachusetts Comprehensive Assessment System (MCAS), a standards-based assessment with 6 years of data, illustrates this issue (Figure 1).[2] In 1998, the first year of MCAS, only 7% of ELL students in Boston Public Schools and 8% of ELL students in Massachusetts overall reached the proficient or above level in 10th-grade English language arts (ELA). Compared to the state level of 38% proficient or above, the gap was approximately 30 percentage points.

In the next few years, perhaps spurred by the addition of the 10th-grade language arts assessment as a graduation requirement (M. Donahue, Boston Public Schools, personal communication, October 21, 2003), Massachusetts 10th-grade ELA scores improved substantially, especially for the state, reaching 61% proficient or above in 2003. However, by 2003, the gap between ELL students and Massachusetts students overall had increased to 49 percentage points—61% for all students versus 12% for ELL students—compared with 30 percentage points in 1998.

In Boston Public Schools, the state's largest district, with approximately 10% ELL students, the ELL gap grew as well, beginning at 11 percentage points in 1998 and increasing to 20 percentage points in 2003. Rapid progress by students overall combined with policies that test ELL students who have lived in the United States for very short periods of time contribute to a growing ELL achievement gap in many states and school districts. In 2003, for example, all ELL students in Massachusetts were required to participate in MCAS testing.

The Massachusetts annual measurable objectives (AMO) trend line (also displayed in Figure 1) shows the difficult challenge that ELL subgroups face in meeting NCLB adequate yearly progress requirements. The annual measurable objective is a target growth rate of students reaching the proficient or above level in a specific subject. In Massachusetts, the target is 48% proficient or above in English language arts for 2003–2004. The target is 56% for 2004–2005, an 8 percentage point increase, and 64% in 2006–2007, another 8 percentage point increase. Unfortunately, ELL statewide performance did not increase 8 percentage points in English language arts during an entire 5-year period (1998–2003), moving unsystematically between 6% and 12% proficient or above. Recognizing again that ELL testing and inclusion policies may contribute to flat ELL English language arts scores, it is clear that Massachusetts ELL students at minimum have a long distance to go to reach the NCLB 100% proficiency rate required by 2014.

Measurement Accuracy

CRESST studies have repeatedly shown that English language learners perform substantially lower on language arts tests compared to mathematics and science tests. Additionally, in our controlled studies measuring accommodations effects, CRESST has been able to repeatedly improve ELL performance by approximately 10-20% on many tests. Modifying, oftentimes simplifying, the language of the test items (see, for example, Abedi &

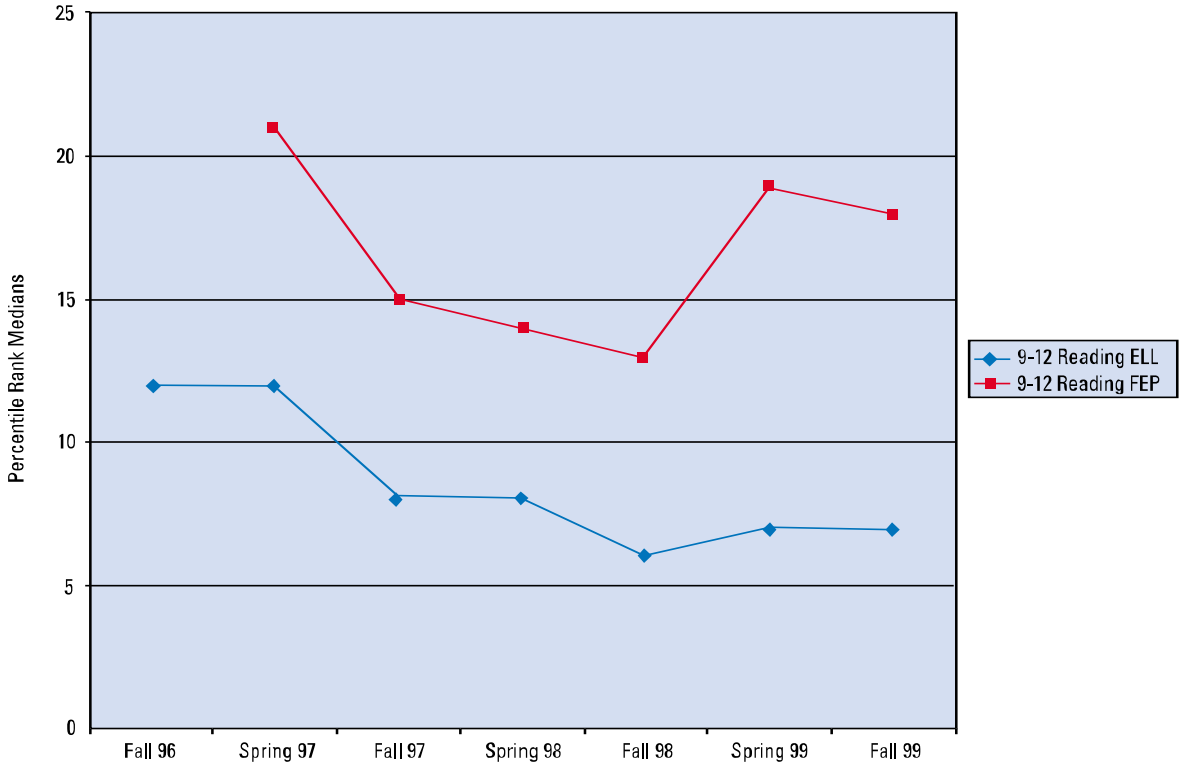

Figure 2. Successive Stanford 9 reading scores for English language learner (ELL) and redesignated Fluent English Proficient (FEP) students in California, Grades 9-12.

Lord, 2001) has consistently resulted in ELL performance improvement without reducing the rigor of the test. These findings suggest to us that low ELL language ability decreases ELL performance on most tests, thus influencing the test as an accurate measure of ELL content knowledge. The test becomes a measure of two skills for the ELL student, subject *and* language.

We note that ELL performance—even when using tests translated into a student's native language or modified-English versions of tests—is still low. Attributing poor performance to the test is inappropriate, but recognizing the downward language effect on assessment outcomes is important. Similarly, in the classroom, the ELL student faces the dual task of learning a school subject *and* a new language.

Instability of the ELL Student Subgroup

Researchers have long postulated that a central cause of flat ELL test scores is the regular removal of high-achieving students from the ELL subgroup. For example, in California, once ELL students become language proficient, they are redesignated as Fluent English Proficient (FEP) and removed from the ELL subgroup. In a study of approximately 14,000 students, we found that redesignation and removal of students from the main ELL group coincided with a modest but significant performance drop by ELL students in reading (Figure 2). Though this is not a claim of causation between ELL redesignation and dropping scores, it reinforces the redesignation dilemma specific to the ELL subgroup.

Recognizing this specific challenge to ELL improvement, the U.S. Department of Education recently approved NCLB plans for several states permitting redesignated English proficient students to remain in the ELL subgroup until they have reached the proficient level on the English language arts section of the state achievement test for several consecutive years (Erpenbach, Forte-Fast, & Potts, 2003). But whether this type of flexibility will provide a long-term solution to the ELL redesignation problem is unknown. Not surprisingly, the high-stakes nature of NCLB has pushed states into submitting Elementary and Secondary Education Act (ESEA) accountability plans that delay consequences to future years, hoping that the next ESEA will be more flexible.

States, districts, and schools with a growing ELL population face an added challenge. A continuous increase of low-achieving ELL students, even if all other factors remain constant, will make it more difficult to achieve adequate yearly progress for both the ELL subgroup and the overall student population.

Factors Outside of a School's Control

CRESST research supports hundreds of other studies showing that nonschool factors, usually parent education level or socioeconomic status, outweigh school factors in their effect on student achievement. In a study of more than 30,000 students, we found these influences even within the ELL population itself. In Figure 3, the gap

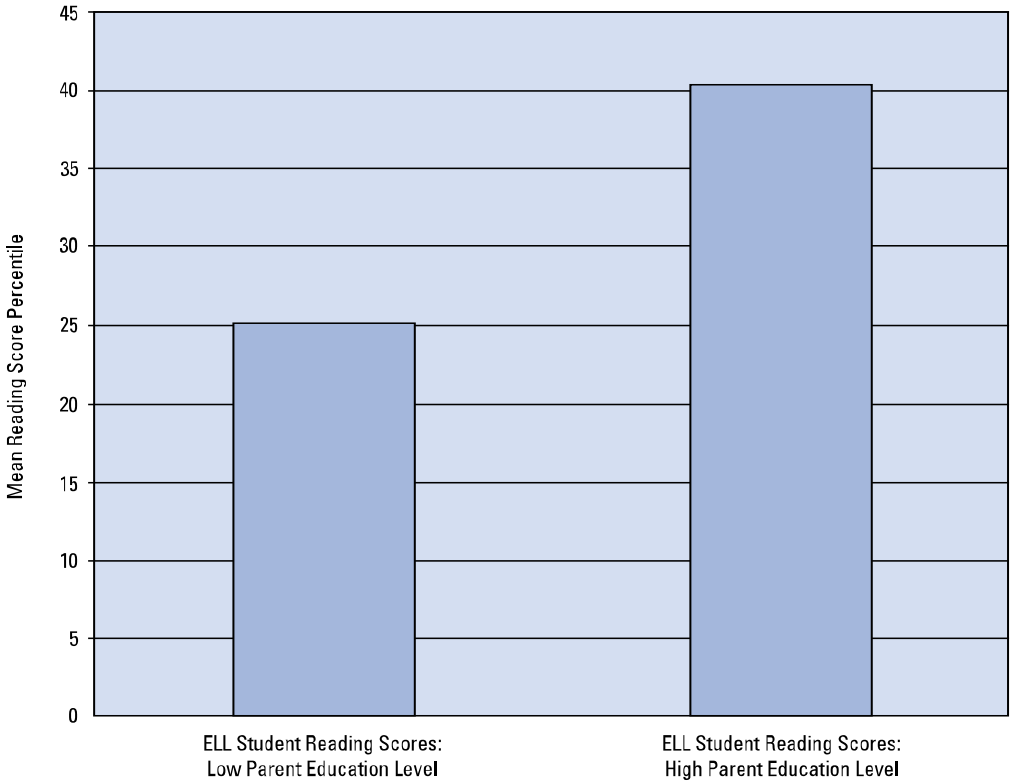

Figure 3. Mean Stanford 9 student reading scores at two parent education levels.

between reading scores of ELL students whose parents had postgraduate education and ELL students whose parents had not graduated from high school was approximately 15 percentile points. Nonschool factors are strong and persistent.

There are a number of other ELL issues that will pose challenges to states, districts, and schools in meeting NCLB goals. These include:

- **ELL identification.** NCLB is helping to create a more common definition of ELL students, but accurate ELL identification remains a challenge. CRESST research has found that many school ELL identification methods produced inconsistent results. Our studies have found a weak relationship between ELL classification codes and language proficiency test scores or achievement test scores. Accurate identification is important in order to provide comparability of ELL performance between states or from district to district or school to school within a state.

- **Subgroup size.** NCLB does not set a minimum number of students necessary to require subgroup reporting but does require large enough numbers to provide statistically reliable data. A low number of students, say 25, will increase the number of schools reporting ELL subgroups but will have the negative effect of more variable performance (Linn, Baker, & Herman, 2002). Setting a higher minimum number of students, say 100 (Hill & DePascale, 2003), will produce less performance variation and thus a more dependable growth rate, but will substantially reduce the number of schools that must report English language learners as a subgroup. Allowing states to have different minimum group sizes will create comparability and fairness issues between state accountability systems.

- **ELL subgroup diversity.** Another challenge to states is the diversity of student achievement within ELL subgroups. ELL performance varies along cultural-linguistic lines, for example, similar to performance differences in the general student population. In one study, we found substantial differences in performance between ELL students with a Chinese-speaking background and ELL students with a Spanish-speaking background. For both the ELL students and the general student group, students with a Chinese-speaking background had significantly higher performance on science and reading tests than students with a Spanish-speaking background.

- **Increased number of subgroups.** Research by Thomas Kane (see Lewis, 2002, pp. 6-7) showed a decreasing probability of schools making specific goals as the number of subgroups increases. It has been our observation that schools with sizeable ELL subgroups tend to have higher numbers of subgroups overall, thereby increasing the probability that those schools will not make adequate yearly progress.

continued on page 5

- **Possible effect of multiple subgroups.** Many ELL students fall into multiple subgroups—for example, special education and cultural-linguistic minority. Consequently, these students' scores frequently are counted multiple times compared with scores of students in the general population and thus may cause a diversion of resources away from those children not in a specific subgroup.

- **NCLB goals and ELL subgroups.** A key NCLB goal is for all subgroups, including ELL students, to reach 100% proficiency in English language arts. However, if ELL students were proficient in English language arts, they would not be ELL students in the first place. Indeed, if NCLB goals were attained, the ELL subgroup would cease to exist.

Notwithstanding these issues and concerns, there are some examples of significant ELL progress. The 2002 National Assessment of Educational Progress (NAEP) state reading results showed positive trends in ELL performance and should give some encouragement to schools, districts, and states. However, comparisons between the 2002 NAEP reading scores and previous results are questionable due to the lower numbers of ELL students participating in earlier NAEP reading assessments and an increasing use of accommodations in recent years (Plisko, 2003). There are state results that show positive signs too, such as the last 3 years of ELL performance on the California Standards Tests. In the following section, we offer suggestions for improving the assessment of ELL students and their achievement.

Recommendations

What can states, districts, and schools do to help their ELL students achieve the rigorous No Child Left Behind goals? We present a few ideas.

- Focus on reading. Though this might seem obvious, CRESST research confirms that ELL students who are better readers, as measured on separate reading tests, perform at higher levels.

- Closely track ELL performance, for both individuals and groups, to identify patterns of improvement or lack of improvement, ideally using multiple measures. To make the substantial gains required by NCLB, schools will need to identify superlative ELL teaching practices and teachers, using that knowledge to help other schools. CRESST's free Quality School Portfolio, for example, allows a school to set performance goals and then measure how well it is doing.

- As mentioned earlier, CRESST research has shown that modifying the language of test questions increases ELL performance, frequently as much as 10-20%. This positive effect is found for content-based tests in different subject areas including mathematics and science. We recommend that all state tests undergo rigorous review for language difficulty and that test questions be modified to reduce the level of unnecessary linguistic complexity and cultural bias.

- NCLB's adequate yearly progress requirements encourage retention of the most capable ELL students within the ELL subgroup, contrary to educational purposes. We believe that ELL redesignation should count for something positive and not serve as a disincentive. Though no perfect solution comes to mind, schools and districts that reach a specific ELL redesignation rate could receive a reduced annual measureable objective or similar incentive that would encourage high redesignation rates. Such concepts would need extensive analysis.

- Encourage test accommodations, but evaluate their validity. CRESST research has shown higher performance when ELL students are provided extra time or a glossary of noncontent-related key terms on the test plus extra time. Because accommodations may overcompensate for a disadvantage, evaluation is important to detect such effects.

- Finally, we believe that performance goals for English language learners, as for all students, should meet what Robert Linn (2003) referred to as "existence proof." For a goal to be within reach of all schools, at least one school should have already attained it. To date we have yet to see a school with a sizeable ELL population that meets the 2014 NCLB requirements.

CRESST Policy Brief 4, *Assessment and Accommodations for English Language Learners*, summer 2001, contains additional information on this topic. Several CRESST research reports listed in the References and Resources section may also be helpful. See CRESST.org for research reports on this and other accountability issues.

References and Resources

Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *Educational Assessment*, *8*, 231-257.

Abedi, J. (2003). *The validity of the classification system for students with limited English proficiency: A criterion-related approach.* Manuscript submitted for

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

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*, 219-234.

Abedi, J., Lord, C., Hofstetter, C., & Baker, E. L. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice*, *19*(3), 16-26.

Abedi, J., Lord, C., & Plummer, J. (1997). *Language background as a variable in NAEP mathematics performance* (CSE Tech. Rep. No. 429). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

California School Board Association. (2003, June 10). *Feds approve California's NCLB plan*. http://www.csba.org/co/top_stories/june03/top_story_6_10_2.htm

Erpenbach, W. J., Forte-Fast, E., & Potts, A. (2003). *Statewide educational accountability under NCLB* (an Accountability Systems and Reporting State Collaborative on Assessment and Student Standards paper). Washington, DC: Council of Chief State School Officers.

Hill, R. (2003). *Options for NCLB designs* [Conference overheads]. Dover, NH: National Center for the Improvement of Educational Assessment. Retrieved 23 November 2003 from http://www.nciea.org/cgi-bin/pubspage.cgi?sorby=pub_date

Hill, R. K., & DePascale, C. A. (2003). *Reliability of No Child Left Behind accountability designs*. Dover, NH: National Center for the Improvement of Educational Assessment.

Lewis, A. (2002, Fall). CRESST 2002 annual conference: "Civil discourse" on assessment. Part 1. *The CRESST Line*, 2, 6-8. Retrieved 20 November 2003 from http://www.cresst.org/products/newsletters_set.htm]

Linn, R. L. (2003, April). *Accountability: Responsibility and reasonable expectations*. Presidential address presented at the annual meeting of the American Educational Research Association, Chicago.

Linn, R. L., Baker, E. L., & Herman, J. L. (2002, Fall). From the directors: Minimum group size for measuring adequate yearly progress. *The CRESST Line*, 1, 4-5. Retrieved 20 November 2003 from http://www.cresst.org/products/newsletters_set.htm

Massachusetts Department of Education. (2003, May 1). *Consolidated state application for state grants under Title IX, Part C, Section 9302 of the Elementary and Secondary Education Act (Public Law 107-110)*.

Plisko, V. W. (2003). *The release of the National Assessment of Educational Progress (NAEP)* The Nation's Report Card: Reading 2002 (Commissioner's Remarks, June 19). Washington, DC: National Center for Education Statistics. Retrieved 19 November 2003 from http://nces.ed.gov/commissioner/remarks2003/6_19_2003.asp

[1]Special thanks to CRESST co-directors Robert Linn and Joan Herman for their valuable review and suggestions for this brief. Our thanks as well to Maryellen Donahue, Boston Public Schools, for her review and comments. Thanks to Seth Leon for his assistance with statistical analysis.

[2]MCAS was selected because it is a long-established standards-based assessment. Our use of MCAS is intended only to illustrate several important points in this brief.

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B960002, as administered by the Institute of Education Sciences, U.S. Department of Education. The findings and opinions expressed in this publication do not reflect the positions or policies of the National Center for Education Research, the Institute of Education Sciences, or the U.S. Department of Education.

GRADUATE SCHOOL OF EDUCATION & INFORMATION STUDIES

UCLA Center for the Study of Evaluation
CSE/CRESST
GSE&IS BLDG MAILBOX 951522
LOS ANGELES CA 90095-1522
ADDRESS SERVICE REQUESTED