

# **Academic Achievement of English Language Learners in Post Proposition 203 Arizona**

**Executive Summary**

**by**

**Wayne E. Wright  
Assistant Professor  
University of Texas, San Antonio**

**Chang Pu  
Doctoral Student  
University of Texas, San Antonio**

Language Policy Research Unit (LPRU)  
Education Policy Studies Laboratory  
College of Education  
Division of Educational Leadership and Policy Studies  
Box 872411  
Arizona State University  
Tempe, AZ 85287-2411

**September 2005**

**EPSL** | EDUCATION POLICY STUDIES LABORATORY  
Language Policy Research Unit

EPSL-0509-103-LPRU

<http://edpolicylab.org>

# **Academic Achievement of English Language Learners in Post Proposition 203 Arizona**

**Wayne E. Wright  
and  
Chang Pu**

University of Texas, San Antonio

## **Executive Summary**

This report reveals the problems with claims made by Arizona state public education officials that English Language Learners (ELLs) are thriving under English-only instruction.

The No Child Left Behind Act of 2001 (NCLB) and the state's accountability system, Arizona LEARNS, require all students, including ELLs, to participate in statewide high-stakes testing. Test scores are the main measure of student achievement under these systems, and labels based on those scores are given to each school (i.e. Highly Performing, Underperforming, etc.). The state education administration's interpretation and strict enforcement of Proposition 203 has ensured that nearly all ELL students in grades K-3 are instructed through the English-only Sheltered English Immersion (SEI) model. They claim that SEI has led to better test scores and increased achievement among ELLs, using as evidence improved test scores and the decrease in the number of schools labeled as "Underperforming." However, analyses of test data for

students in grades two through five and changes in the state accountability system revealed the contrary; they exposed serious achievement gaps between ELLs and their counterparts, and proved that positive looking improvements in school accountability labels mask test-score decline in a large number of elementary schools.

From 2002 to 2004, students in Arizona were required to take two standardized tests: Arizona Instrument to Measure Standards (AIMS), a test given in grades three, five, eight, and high school that is designed to measure student achievement against state standards; and the Stanford Achievement Test *Ninth Edition* (Stanford 9), a test given in grades two through nine that is designed to measure student achievement against the national average. The state has divided test score data into two categories: ALL (Category 1) and ELL (Category 2). The labels are misleading: The ALL category excludes the scores of ELL students who have been enrolled in public school for less than four years, thereby excluding the scores of the ELL students with the lowest levels of English proficiency. The report's analyses focus mostly on third grade AIMS test scores and the Stanford 9 test scores of elementary school students as they progressed from one grade to the next between 2002 and 2004. The key findings are:

- The overwhelming majority of third grade ELLs fail the AIMS test in contrast to ALL students, and ELLs score well below the 50th percentile on the Stanford 9 and well below students in the ALL category.
- There is a general pattern of higher test scores on AIMS in 2003, followed by decline in 2004 for both ALL and ELL students on the Reading and Math subtests.

- ELL student percentile rankings on the Stanford 9 rose slightly in 2003 followed by a decline in 2004 while ALL student rankings remained relatively stable.
- Improvement in test scores in 2003 corresponds with a period of greater flexibility for schools in offering ESL and bilingual education, while the decline of scores in 2004 corresponds to a period of strict enforcement of Proposition 203 and mandates for English-only instruction.
- The sudden increase in 2004 of ELLs passing the AIMS Writing subtest is questionable, as there was decline or no significant growth on all other subtests for both the AIMS and Stanford 9, and as similar gains were not evident for ALL students.
- In terms of the percent passing the AIMS test, ELL students trailed behind ALL students by an average of 33 percentage points in Math, 40 points in Reading, and 30 points in Writing.
- On the Stanford 9, ELL students trailed behind ALL students by an average of 28 percentile points in Language, 26 points in Math, and 33 points in Reading. The gap increased for all Stanford 9 subtests between 2003 and 2004.
- The narrowing of the achievement gap in AIMS Reading and Math is actually a function of ALL student scores decreasing at a higher rate than decreases in ELL scores.

- ALL students score lower on the AIMS and Stanford 9 in ELL-Impacted elementary schools (schools that test 30 or more ELL students in third grade) than they do in other elementary schools.
- Lack of reliable data: There are discrepancies in the number of ALL and ELL students tested on the AIMS and Stanford 9 within each year and across the three years that are inconsistent with the rapidly growing student population of Arizona. This raises questions on whether some student scores are missing from the data reported to the public, or if students were systematically excluded from taking specific tests.

This report also analyzes the changes in school labels under Arizona LEARNNS and NCLB between 2002 and 2004. In 2002, the Arizona LEARNNS labels were: Excelling, Maintaining, Improving, and Underperforming. In 2003, the labels were changed to: Excelling, Highly Performing, Performing, Underperforming, and Failing. These labels are based primarily on the test performance of students in the ALL category, which excludes most ELL scores. An analysis of the numbers of schools in each category throughout this time period along with the test data for the corresponding years revealed the following:

- There were increases in the number of “Performing” and “Excelling” schools in 2004 despite the general trend of flat or declining AIMS and Stanford 9 scores.
- Arizona LEARNNS labels and NCLB AYP designations are not reflective of a school’s success (or lack thereof) with ELL students as these labels and

designations are based on ALL score data which excludes most ELL test scores.

- Improvements in Arizona LEARNS labels and NCLB's AYP designations are masking the harm that current state language and testing policies are having on ELL students.

Close monitoring of ELL test scores is needed by policy makers and relevant stakeholders. A system is also needed for mutually exclusive categories of ELL and non-ELL students, and mechanisms are needed to track the progress of ELL students even after they are redesignated as fluent English proficient. State policy makers are encouraged to reconsider the narrow requirements and current strict enforcement of Proposition 203. In addition, rather than forcing ELLs to take English-only high-stakes tests only to exclude many of their scores from state and federal accountability formulas, state policy makers are encouraged to advocate for changes in the requirements of NCLB, or at the very least, heed the federal law's requirement to test ELLs in the *language* and *form* most likely to yield valid and reliable information about what students know and can do.