Education and the Public Interest Center
University of Colorado at Boulder

EPRU Education Policy Research Unit
Arizona State University

DOCUMENT REVIEWED:	"Has Progress Been Made in Raising Achievement for English Language Learners?"
AUTHORS:	Naomi Chudowsky and Vicktor Chudowsky
PUBLISHER/THINK TANK:	Center on Education Policy
DOCUMENT RELEASE DATE:	April 2010
REVIEW DATE:	May 19, 2010
REVIEWER:	Jeff MacSwan, Arizona State University
E-MAIL ADDRESS:	macswan@asu.edu
PHONE NUMBER:	(480) 965-4967

SUGGESTED CITATION:
MacSwan, J. (2010). *Review of "Has Progress Been Made in Raising Achievement for English Language Learners?"* Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. Retrieved [date] from http://epicpolicy.org/thinktank/review-progress-ELL

Summary of Review

The Center on Education Policy (CEP) report, *Has Progress Been Made in Raising Achievement for English Language Learners?*, finds that some states have seen increases in the number of English language learners (ELLs) meeting proficiency standards under No Child Left Behind (NCLB), while others have seen decreases. The report notes some limitations in the data it uses. The CEP report, however, has some specific weaknesses in its research methods that undermine its findings. The CEP report seriously underestimates the significance of language of instruction as a source of error in ELL achievement test scores. Further, it errs in implying that its findings justify an inference of a causal relationship between observed changes in percentages of ELLs meeting achievement benchmarks and improvements in academic achievement for ELLs. Given the limitations in the data, it is inappropriate to draw conclusions from the data summarized in the CEP report.

Review

I. Introduction

Has Progress Been Made in Raising Achievement for English Language Learners?,[1] a report published by the Center on Education Policy (CEP), focuses on the achievement of English language learners (ELLs) under current federal NCLB policy. The report examines changes in the percentages of ELLs identified as meeting the various levels of proficiency in data reported by states from 2006 to 2008.

II. Findings and Conclusions of the Report

The report reaches eight specific conclusions (pp. 1,2). Two conclusions focus on data quality:

- "Because of language barriers, test results for English language learners may not accurately reflect what these students know and can do. …
- "The number of ELL test-takers has changed rapidly in many states, which complicates efforts to track achievement trends for this group. …

Six conclusions focus on ELL achievement:

- "In grade 4 reading and math, rising percentages of English language learners have reached three achievement levels—basic, proficient, and advanced—although gains are less prevalent at the advanced level in reading. …
- "English language learners have made progress in reaching state proficiency benchmarks in reading and math at all three grade levels analyzed—elementary, middle, and high school—although gains are less prevalent in high school than at the other grade levels. …
- "Trends for English language learners are mostly positive in states with the largest number of ELL test-takers. In the 10 states with the most ELL test-takers, positive trends outnumbered negative trends. …
- "English language learners in grade 4 are doing better in math than in reading at all three achievement levels. …
- "Very large differences in percentages proficient exist between English language learners and other students. …
- "Percentages proficient for English language learners vary widely by state, more so than for students who are not ELLs. …"

III. The Report's Rationale for Its Findings and Conclusions

The report's authors indicate that their "main unit of analysis is the trend line, which is a record of change from 2006 through 2008 in the performance of the ELL subgroup …" for all the states in the aggregate. This unit of analysis, they believe, will help them with their primary task, namely, "to determine whether ELLs have made progress across the achievement spectrum" (p. 4). While they present some cautions, the authors believe that the data support the general conclusion that states have made progress in the raising achievement of ELLs. The authors present descriptive summaries of student scores. They do not use any statistical tests to compare student outcomes.

IV. The Report's Use of Research Literature

The report makes limited use of relevant research literature. Some attention is given to research related to special accommodations for ELL test-takers and to validity concerns

about ELL test scores. The report does not, however, take advantage of methodological insights of related studies published elsewhere.[2]

V. Review of the Report's Methods

There are significant concerns related to the quality and suitability of the data that the report uses to support its conclusions. The report's authors note some of these limitations, but underestimate their significance, and they advance conclusions based on the data even though the limitations they note should prevent such conclusions from being offered.

Language of the Test as a Source of Error

The report notes that ELLs' test scores in English may "not accurately reflect what these students know" and notes that "questions remain about the reliability of test scores for ELLs and the validity of inference drawn from these scores" (p. 1). However, this assessment severely understates the problem associated with using ELLs' scores in English; more accurately, one might characterize the scores as misleading and inappropriate for students with limited English proficiency.

Language proficiency in general has been shown to influence performance on achievement tests.[3] Pilkington, Piersel, and Ponterotto (1988) reported that the home language of a child influenced the predictive validity of kindergarten achievement measures.[4] These studies suggest language proficiency plays a role in young children's performance on achievement tests. The *Standards for Educational and Psychological Testing* warns that when testing a non-native speaker in English, the test results cannot be expected to reflect accurately the abilities and competencies being measured if test performance depends on the test takers' knowledge of English.[5] Because states' academic proficiency tests are designed to measure knowledge of content standards and are specifically not designed in relation to a theory of language proficiency, limited English proficiency must be regarded as a major source of error in the measurement of academic achievement.

To illustrate the problem associated with relying on assessment data from ELLs tested in English, imagine a test of academic achievement administered to a group of English-speaking elementary school students in Spanish. Because relatively few of the students know Spanish, they would have considerable difficulty understanding the questions. We would expect their limited Spanish ability to negatively affect their test scores. Furthermore, because the test claims to measure academic achievement and not Spanish, we would regard these circumstances as a source of measurement error. The scores would be of little or no value, as they would not reflect students' subject matter knowledge.

We know with some certainty that all the students in the ELL subgroup have been assessed as having limited English proficiency by a state language proficiency test. We can therefore conclude that the achievement test scores summaries in the CEP report are not valid, as the students have only limited ability to understand the language of the academic achievement test.

Changing Student Cohort Composition Makes Comparisons Inappropriate

Although the reports' authors are clearly aware that the data relate to changing percentages of ELLs and to states' relative success in increasing the percentages of ELLs meeting proficiency benchmarks, they sometimes state conclusions in terms of student-level growth.

For example, the report asserts that the data show that "English language learners have made progress in reaching state proficiency benchmarks in reading and math at all three grade levels …" (p. 2). However, the data do not permit conclusions regarding year-to-year changes in ELL performance because student cohorts are not followed over time in the study. The composition of the ELL subgroup changes dramatically from year to year because students are exited from the subgroup each year based on achievement score benchmarks ("skimming effect")[6]; because ELLs and other lower income students are highly mobile, moving across districts, states, and national boundaries; and because state policy fluctuates with regard to conditions under which ELL scores are included in data summaries. These facts are also suggested by the report's observation that numbers of test-takers changed dramatically over time in some states. Given these considerations, it is not only unlikely that the ELL subgroup is made up of the same students each year, but it is virtually certain that it is substantially different each year. Hence, drawing growth-related conclusions from data with these dramatic limitations is inappropriate.

Causal Inference is Not Valid

The report does not focus on changes in data reflecting student achievement outcomes (test score data, for instance), but rather on the percentage of students meeting state criteria each year. The question asked in its title—"Has progress been made in raising achievement for English language learners?"—is not actually addressed in the report and cannot be addressed using these data or research methods. The question asks us to infer a direct causal relationship between changes in reported percentages of ELLs meeting year benchmarks and improvements in academic achievement for ELLs. However, no methodological approach to causal inference is followed in the report, and multiple competing explanations for the annual changes, each as plausible as the next, are not addressed. Consider, for instance, the following possible alternative explanations for the change in students' test scores:

- Standardized test scores tend to increase over time regardless of interventions.[7]
- Schools have focused more strategically on the sample of items likely to be addressed in the assessments.
- Districts are postponing reclassification of high-scoring ELLs.
- Lower scoring ELLs are relocating to states reporting declines in the percentage of students meeting proficiency.
- States with higher percentages of ELLs meeting proficiency have easier tests.
- States with higher percentages of ELLs meeting proficiency have negotiated with the US Department of Education to postpone reporting scores of newcomers.

A causal inference would require, at a minimum, that competing explanations be eliminated. While the report notes these potential complications, regrettably these concerns do not prevent the authors from drawing specific conclusions about the achievement trajectories of ELLs.

VI. REVIEW OF THE VALIDITY OF THE FINDINGS AND CONCLUSIONS

The conclusions about the academic performance of ELLs drawn in the report are not supported by the evidence presented for three specific reasons. The report underestimates the significance of limited English proficiency as a source of error in ELL test scores. Data are inappropriately presented as achievement trends for ELL students even though the composition of student cohorts is different each year. Finally, the report's authors reason that changes in percentages of students reaching

state benchmarks reflect actual changes in student achievement over time, even though competing explanations are not addressed.

Given these limitations, it is inappropriate to draw conclusions of any kind from the data; therefore it is not only not useful, it is misleading to present summaries of the data, because readers will tend to draw conclusions based on the data summaries alone.[8]

VII. USEFULNESS OF THE REPORT FOR GUIDANCE OF POLICY AND PRACTICE

The report is not a useful guide for policy or practice related to ELLs. The data analyses presented do not provide useful insight into ELL progress, and no judgments can be made regarding the relative success of current federal policy based on the report.

Notes and References

[1] Churdowsky, N. & Chudowsky, V. (2010). *Has Progress Been Made in Raising Achievement for English Language Learners?* Washington, DC: Center on Education Policy.

[2] Parrish, T. B., Merickel, A., Pérez, M., Linquanti, R., Socias, M., Spain, A., Speroni C., Esra, P., Brock, L., & Delancey, D. (2006). *Effects of the Implementation of Proposition 227 on the Education of English Learners, K–12: Findings from a Five-Year Evaluation*. San Francisco: WestEd.

Thompson, M., DiCerbo, K., Mahoney, K., & MacSwan, J. (2002). Exito en California? A validity critique of language program evaluations and analyses of English Learner Test Scores. *Education Policy Analysis Archives*, 10(7). Retrieved May 9, 2010, from http://epaa.asu.edu/epaa/v10n7/

[3] Ulibarri, D. M., Spencer, M. L., & Rivas, G. A. (1981). Language proficiency and academic achievement: A study of language proficiency tests and their relationship to school ratings as predictors of academic achievement. *NABE Journal, 5*, 47-80.

[4] Pilkington, C. L., Piersel, W. C., & Ponterotto, J. G. (1988). Home language as a predictor of first-grade achievement for Anglo- and Mexican-American children. *Contemporary Educational Psychology, 13*, 1-14.

[5]American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

[6] Thompson, M., DiCerbo, K., Mahoney, K., & MacSwan, J. (2002). Exito en California? A validity critique of language program evaluations and analyses of English Learner Test Scores. *Education Policy Analysis Archives*, 10(7). Retrieved May 9, 2010, from http://epaa.asu.edu/epaa/v10n7/

[7] Linn, R., Graue, E., & Sanders, N. (1990). Comparing state and district test results to national norms: The validity of claims that "everyone is above average." *Educational Measurement: Issues and Practice, 9*, 3, 5-14.

[8] For instance, the Lexington Institute released a press release regarding the CEP report, claiming, "New research on state test scores by the Center on Education Policy confirms … that changes made in federal education policies for English learners as part of the No Child Left Behind Act have produced substantially higher test scores overall for this critical, growing student population."

Soifer, Don. (May 4, 2010). Results Show Inclusion, Exclusion Impact English Learners. Arlington, VA: Lexington Institute. Retrieved May 16, 2010, from http://www.lexingtoninstitute.org/results-show-inclusion-exclusion-impact-english-learners?a=1&c=1136.

The Think Tank Review Project is made possible by funding from the Great Lakes Center for Education Research and Practice.