



Jamal Abedi

Dr. Abedi's research interests include psychometrics and test and scale development. His recent work includes validity studies for the National Assessment of Educational Progress (NAEP) focusing on the impact of language background on students' performance and the dimensionality of NAEP math subscales.

Assessment *and* Accommodations for English Language Learners: *Issues and Recommendations*

Jamal Abedi

Today, one out of seven children in the U.S. speaks a language other than English at home.¹ In California, Florida, New Mexico, and Texas, English language learners (ELLs) exceed 10% of the student population.² This policy brief addresses the inclusion of ELL students³ in large-scale assessments and ELL assessment accommodations.

Both federal and state legislation now require inclusion of all students, including English language learners (ELLs), in large-scale assessments⁴ based on a number of key factors:

- Inclusion provides a more accurate picture of overall student achievement and growth.
- Inclusion makes individual diagnostic information available to parents of ELL students, their teachers, and school administrators.
- Inclusion can provide evidence that ELL students have reached proficiency and therefore should no longer be considered English language learners.
- Inclusion may allow for specific policies and funding to improve the performance of ELL students.

CRESST partner Lorrie Shepard and colleagues point out that “inclusion signals the commitment of the educational system to support the academic progress of all its students; and it ensures the representativeness of the data reported.”⁵ However, inclusion of ELL students creates specific accountability policy challenges:

1. Which ELL students should be included in large-scale assessments is highly variable from state to state as well as between states and other major assessments such as the National Assessment of Education Progress (NAEP).
2. States differ on inclusion/exclusion policies, type and use of accommodations, reporting of accommodated assessments, and related issues.⁶
3. National norm-referenced standardized tests may have been normed using mostly native English speakers. Accuracy of results may be affected for states with large ELL student populations.
4. If ELL performance is low, we may not know whether the cause is due to limited language skills, low content knowledge, or a combination of both.⁷
5. Inclusion does not address different language and knowledge skills within the ELL population. Although ELL students as a group usually underperform compared to non-ELL students,⁸ their range of performance is quite broad.⁹
6. ELL test results may be used inappropriately, either to sort or retain students, or to pull children out of their regular classrooms and into less than ideal instructional programs.

Accommodations Issues You Should Know

The complex ELL assessment environment is made even more challenging with the introduction of accommodations for ELL students. (See Butler and Stevens¹⁰ for a comprehensive list of frequently used accommodations.) Accommodations are intended to level the playing field, that is, to make language less of a factor, or ideally a non-factor, when measuring performance. However, the use of accommodations requires a complex set of practical and technical decisions, and the research in support of these choices is thin. Such decisions should be informed by the following:

- **Validity:** Does provision of accommodation alter the construct of the assessment?
- **Effectiveness:** What accommodation strategies would be the most effective in reducing performance gaps between ELL students and non-ELL students that are due to language factors?
- **Differential impact:** Which student background characteristics impact accommodated assessment?
- **Feasibility:** Which accommodation(s) are more feasible, particularly in large-scale assessments?

Among these guidelines, the most important is validity. Accommodations should reduce the impact of language but not give LEP students an “unfair advantage” over students not receiving accommodations.¹¹ However, CRESST research using randomized assignment methods suggests that accommodations may threaten the validity of an assessment, either through over-accommodation or by significantly changing the standardized administration conditions under which the assessment was developed.

Abedi and colleagues¹² found that non-ELL students benefited more from certain forms of accommodation than ELL students did. Dan Koretz¹³ found that accommodations provided in Kentucky produced unrealistically high scores for some special needs students. See Abedi et al.¹⁴ for a discussion of challenges to the validity of an assessment through accommodations.

What CRESST Research Tells Us

- Language proficiency strongly relates to test performance. Students designated ELL by their schools score significantly lower than non-ELL students on many science and math questions.¹⁵ However, the performance gap decreases, or even disappears, on math items that have relatively low language demands, such as math computation.¹⁶
- ELL students who are better readers, as measured by separate reading tests, perform better on questions with high language demands.¹⁷
- Translating test items from English to a student’s native language does not significantly improve ELL performance when the language of instruction is not the student’s native language.¹⁸
- **The only accommodation that narrowed the gap between ELL and non-ELL students was linguistic modification of those test questions with excessive language demands.**¹⁹
- In addition to language proficiency, other background factors influence ELL performance. These factors include length of time in the United States, overall grades, and student mobility.²⁰
- Many accommodations require substantial amounts of additional administrative time that may increase costs substantially.

Recommendations

CRESST research, supported by other research we have reviewed, leads us to make the following recommendations to policymakers and educators involved in the assessment of ELL students. As we accumulate ongoing research, these recommendations will expand.

1. National clarification is needed on a common definition of ELL students. Even more important are criteria for ELL classifications and appropriate accommodations. Comparability between states, as required in current proposals to reauthorize the Elementary and Secondary Education Act, will otherwise be impossible.
2. Translating test items from English to other languages may not be a successful accommodation when ELL students are taught in English. Our data suggest that the language of the assessment should match the student's primary language of instruction.
3. Student background variables, including language background, are strong indicators of preparedness for participation in large-scale assessments. We recommend that states and other large-scale assessments endeavor to collect background information including length of time living in the United States, type and amount of language spoken in the home, proficiency level in English and student's native language, and number of years taught in both languages.
4. Another suitable readiness indicator is a student's proficiency in academic English. An additional benefit of an external language proficiency measure is that it can be used to suggest appropriate accommodations and monitor language progress over time.
5. Modify test questions to reduce unnecessary language complexity during the development and improvement of all large-scale assessment programs. Reducing language complexity helps to narrow the performance gap between native English speaking students and ELL students.
6. Customized dictionaries are a viable alternative to providing traditional dictionaries as accommodations.²² A traditional dictionary may provide ELL students an unfair advantage on certain types of tests.
7. Feasibility considerations are important. National and state assessments involve a large number of ELL students, so accommodations have substantial costs. Providing dictionaries or glossaries to all ELL students, administering assessments one-on-one to students, or reducing the language complexity of test items may exceed a school, district, or state's capability. Cost-benefit analyses should be considered. At minimum, accommodations' costs should be tracked and evaluated.
8. Intended and unintended accommodations effects must be monitored and evaluated closely. Ideally, accommodations will have no effect on native English speaking students, while reducing the language barrier for ELL students.²³ With states increasingly moving to reward or sanction schools based on test results, evaluating accommodations effects takes on added, schoolwide importance.

Key CRESST Research Finding

Some accommodations are more effective than others. Providing extra time, a glossary of key terms on the test *plus* extra time, or reducing the language complexity of the test questions resulted in substantially higher test scores for ELL and non-ELL students. *Providing a glossary without extra time did not increase ELL performance, possibly due to information overload.*²¹

Recommendations (continued)

9. While we dislike concluding that more research is needed, the complexity of accommodations and their relative newness to education accountability programs require additional investigation. For example, we need to know whether accommodations have different effects by background factors or by subgroups (e.g., Spanish, Vietnamese, and Cambodian). Other areas of research we need to investigate are the effects of multiple accommodations and why some accommodations work better at certain grade levels than others. We also need to improve our reporting of assessments when accommodations are provided.

Contact Jamal Abedi at jabedi@cse.ucla.edu

A complete set of reference notes is available on the CRESST Web site at:
<http://www.cse.ucla.edu/CRESST/Newsletters/polbrf4ref.pdf>

www.cse.ucla.edu

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number **R305B960002:01**, as administered by the Office of Educational Research and Improvement, U.S. Department of Education. The findings and opinions expressed in this publication do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement or the U.S. Department of Education. To order copies of this policy brief, contact Kim Hurst, 310-794-9140, email: kim@cse.ucla.edu, or write to Kim at CRESST/UCLA, GSE&IS Building, Mailbox 951522, Los Angeles, CA 90095-1522.

UCLA Center for the Study of Evaluation
CSE/CRESST
GSE&IS BLDG MAILBOX 951522
LOS ANGELES CA 90095-1522
ADDRESS SERVICE REQUESTED
EEX1

NON PROFIT ORG.
U.S. POSTAGE
PAID
U.C.L.A.

Notes

1. Garcia, E. (1998). *Excellence and equity for language minority students: Critical issues and promising practices*. Chevy Chase, MD: The Mid-Atlantic Equity Center.
2. Mazzeo, J., Carlson, J. E., Voelkl, K. E., & Lutkus, A. D. (2000). *Increasing the participation of special needs students in NAEP: A report on 1996 NAEP research activities* (NCES 2000-473). Washington, DC: U.S. Department of Education, National Center for Education Statistics. [Note: The percent is based on data for public schools, Grades 4 and 8.]
3. The term "English language learner" (ELL; see LaCelle-Peterson, M., & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review*, 64, 55-75) refers to students who are not native speakers of English and are not as proficient in English as the native speakers are. A subgroup of these students with a lower level of English proficiency is referred to as "limited English proficient" (LEP). The term LEP is used primarily by government-funded programs to classify students, as well as by the National Assessment of Educational Progress (NAEP) for determining inclusion criteria. In this policy brief we use English language learner (ELL) to refer to students who are not native English speakers and who are not reclassified as fluent in English.
4. Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice*, 19 (3), 16-26.
See also Note 2, Mazzeo, Carlson, Voelkl, & Lutkus (2000).
5. Shepard, L., Taylor, G., & Betebenner, D. (1998). *Inclusion of limited-English-proficient students in Rhode Island's grade 4 mathematics performance assessment* (CSE Tech. Rep. No. 486). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
6. Rivera, C., Stansfield, C. W., Scialdone, L., & Sharkey, M. (2000). *An analysis of state policies for the inclusion and accommodation of English language learners in state assessment programs during 1998-1999*. Arlington, VA: The George Washington University, Center for Equity and Excellence in Education.
7. Abedi, J., Leon, S., & Mirocha, J. (2001). *Impact of students' language background on standardized achievement test results: Analyses of extant data*. Los Angeles: University of California, 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. (1998). *Impact of selected background variables on students' NAEP math performance* (CSE Tech. Rep. No. 478). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
Abedi, J., Lord, C., & Plummer, J. R. (1997). *Final report of 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.
American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
Garcia, G. E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. *Reading Research Quarterly*, 26, 371-391.
LaCelle-Peterson, M., & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review*, 63, 55-75.
Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23 (1), 13-23.
See also Note 4, Abedi, Lord, Hofstetter, & Baker (2000).

8. See Note 7, Abedi & Lord (2001).
 9. Butler, F. A., & Castellon-Wellington, M. (2000). Students' concurrent performance on tests of English language proficiency and academic achievement. In *Validity of administering large-scale content assessments to English language learners: An investigation from three perspectives* (Final Deliverable to OERI/OBEMLA; pp. 51-83). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
 10. Butler, F. A., & Stevens, R. (1997). *Accommodation strategies for English language learners on large-scale assessments: Student characteristics and other considerations*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
 11. Thurlow, M., Liu, K., Erickson, R., Spicuzza, R., & El Sawaf, H. (1996). *Accommodations for students with limited English proficiency: Analysis of guidelines from states with graduation exams* (Minnesota Report No. 6). Minneapolis: University of Minnesota, National Center on Educational Outcomes.
 12. See Note 4, Abedi, Lord, Hofstetter, & Baker (2000).
 13. Koretz, D. (1997). *The assessment of students with disabilities in Kentucky* (CSE Tech. Rep. No. 431). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
 14. See Note 4, Abedi, Lord, Hofstetter, & Baker (2000).
 15. Abedi, J., & Leon, S. (1999). *Impact of students' language background on content-based performance: Analyses of extant data*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
 16. See Note 7, Abedi, Leon, & Mirocha (2001).
 17. See Note 15, Abedi & Leon (1999), and Note 7, Abedi, Leon, & Mirocha (2001).
 18. See Note 7, Abedi, Lord, & Hofstetter (1998).
 19. See Note 4, Abedi, Lord, Hofstetter, & Baker (2000).
 20. See Note 4, Abedi, Lord, Hofstetter, & Baker (2000), and Note 7, Abedi, Lord, & Hofstetter (1998).
 21. See Note 4, Abedi, Lord, Hofstetter, & Baker (2000).
 22. Abedi, J., Courtney, M., Mirocha, J., Leon, S., & Goldberg, J. (2000). *Language accommodation for large-scale assessment in science*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Abedi, J., Lord, C., Kim, C., & Miyoshi, J. (2000). *The effects of accommodations on the assessment of LEP students in NAEP*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.