

REL Pacific Ask A REL Response

English Learners

June 2018

Question:

What are the effects on student performance when testing occurs in a student’s primary vs. secondary language?

Response:

Following an established REL Pacific research protocol, we conducted a web-based search for resources related to differences in outcomes when students are tested in their primary versus their secondary language (see Methods section for search terms and resource selection criteria). We searched for information specific to the Pacific region, but also expanded the search to look for similar language issues around the world. The compiled sources have been organized into the following categories:

General Information on the Assessment of English Language Learners
Dual Language and English as a Foreign Language

References are listed in alphabetical order, not necessarily in order of relevance. Descriptions of the resources are quoted directly from the publication abstracts. We have not evaluated the quality of references and the resources provided in this response. We offer them only for your reference. Also, we searched the references in this response from the most commonly used research resources, but they are not comprehensive and other relevant references and resources may exist.

Research References

General Information on the Assessment of English Language Learners

Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *Educational assessment*, 8(3), 231–257.

https://www.tandfonline.com/doi/pdf/10.1207/S15326977EA0803_02

From the abstract: “Using existing data from several locations across the U.S., this study examined the impact of students’ language background on the outcome of achievement tests. The results of the analyses indicated that students’ assessment results might be confounded by their language background variables. English language learners (ELLs) generally perform

lower than non-ELL students on reading, science, and math—a strong indication of the impact of English language proficiency on assessment. Moreover, the level of impact of language proficiency on assessment of ELL students is greater in the content areas with higher language demand. For example, analyses showed that ELL and non-ELL students had the greatest performance differences in the language-related subscales of tests in areas such as reading. The gap between the performance of ELL and non-ELL students was smaller in science and virtually nonexistent in the math computation subscale, where language presumably has the least impact on item comprehension.

The results of our analyses also indicated that test item responses by ELL students, particularly ELL students at the lower end of the English proficiency spectrum, suffered from low reliability. That is, the language background of students may add another dimension to the assessment outcome that may be a source of measurement error in the assessment for English language learners.

Further, the correlation between standardized achievement test scores and external criterion measures was significantly larger for the non-ELL students than for the ELL students. Analyses of the structural relationships between individual items and between items and the total test scores showed a major difference between ELL and non-ELL students. Structural models for ELL students demonstrated lower statistical fit. The factor loadings were generally lower for ELL students, and the correlations between the latent content-based variables were also weaker for them.

We speculate that language factors may be a source of construct-irrelevant variance in standardized achievement tests (Messick, 1994) and may affect their construct validity.”

Abella, R., Urrutia, J., & Shneyderman, A. (2005). An examination of the validity of English-language achievement test scores in an English language learner population. *Bilingual Research Journal*, 29(1), 127–144. <https://eric.ed.gov/?id=EJ724701>

From the abstract: “Approximately 1,700 English language learners (ELLs) and former ELL students, in Grades 4 and 10, were tested using both an English-language (Stanford Achievement Test, 9th ed.) and a Spanish-language (Aprenda, 2nd ed.) achievement test. Their performances on the two tests were contrasted. The results showed that ELL students, for the most part, answered more items correctly on a home-language mathematics test, compared to a similar English language math test, regardless of their level of home-language literacy. Additionally, former ELL students are often unable to exhibit their content-area knowledge on English-language achievement tests, possibly due to language and cultural barriers. In summary, the results show that the achievement test results of ELL students, when tested in English, are not always valid measures of their content-area knowledge.”

Fox, J., & Cheng, L. (2007). Did we take the same test? Differing accounts of the Ontario Secondary School Literacy Test by first and second language test-takers. *Assessment in Education*, 14(1), 9–26. <https://eric.ed.gov/?id=EJ763922>

From the abstract: “Within the context of increasing numbers of second language (L2) learners in Canadian schools and expanding standards-driven testing frameworks, a passing score on the Ontario Secondary School Literacy Test (OSSLT) is a recently imposed secondary school graduation requirement in Ontario. There is evidence, however, that tests designed on the basis of first language (L1) populations may have lower reliability and validity for L2 students. This study elicited accounts of the OSSLT in 33 focus groups of 22 L1 students and 136 L2 students, attending 7 Ontario secondary schools, prior to and immediately after the March 2006 test administration. The results suggest important differences in L1 and L2 accounts of test constructs and suggest a gap between what is valued as literacy on the test and what is valued in classroom literacy practice, raising some concern regarding the test’s consequential validity. By examining how different groups of test-takers interpret test constructs and the interaction between these interpretations, test design, and accounts of classroom practice, we may better address issues of fidelity in test construct representation (i.e., understand what may constitute construct under-representation and construct-irrelevant variance). This study highlights what may make a test more L2-friendly, i.e. what supports (or impedes) L2 test performance. Although in the washback literature test-taker accounts of tests have been the least researched, the results of this study suggest that such accounts have the potential to increase test fairness, enhance the validity of inferences drawn from test performance, improve the effectiveness of accommodation strategies, and promote positive washback.”

Scheffel, D., Lefly, D., & Houser, J. (2012). The predictive utility of DIBELS reading assessment for reading comprehension among third grade English language learners and English speaking children. *Reading Improvement*, 49(3), 75–92.
<https://eric.ed.gov/?id=EJ986934>

From the abstract: “The study addresses the extent to which subtests on the Dynamic Indicators of Basic Early Literacy Skills Reading Assessment (DIBELS; Good & Kaminski, 2002) predict student success on a measure of reading comprehension and if prediction is consistent for native and second English Language Learners. 2,649 elementary students were assessed on a reading comprehension measure, of which 29.7% were English Language Learners. Descriptive and analytic statistics were generated including bivariate correlation analysis split by language proficiency. Critical measures and suggested cutoff values (Good, Simmons, et al., 2002) were evaluated for predictive utility by visualization of Receiver Operating Characteristic (ROC) curves (Swets, Dawes, & Monahan, 2000), and comparison of the area-under-the-curve (AUC) values. DIBELS better predicts children who are at "low risk" than those "at risk;" however, DIBELS correctly classifies children "at risk" better for ELL than non-ELL students in third grade.”

Stevens, R. A., Butler, F. A., & Castellon-Wellington, M. (2001). *Academic language and content assessment: Measuring the progress of English language learners (ELLs)*. Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student

Testing, Graduate School of Education & Information Studies, University of California, Los Angeles. NO ABSTRACT

Dual Language and English as a Foreign Language

Guzman-Orth, D., Lopez, A. A., & Tolentino, F. (2017). A Framework for the Dual Language Assessment of Young Dual Language Learners in the United States. *ETS Research Report Series*. <https://eric.ed.gov/?id=EJ1168394>

From the abstract: Dual language learners (DLLs) and the various educational programs that serve them are increasing in number across the country. This framework lays out a conceptual approach for dual language assessment tasks designed to measure the language and literacy skills of young DLLs entering kindergarten in the United States. Although our examples focus on Spanish-English DLLs, we anticipate that our recommendations could be broadly applied to other language combinations with appropriate adaptations for each language.

Lindholm-Leary, K., & Hernández, A. (2011). Achievement and language proficiency of Latino students in dual language programmes: Native English speakers, fluent English/previous ELLs, and current ELLs. *Journal of Multilingual and Multicultural Development*, 32(6), 531–545. <https://eric.ed.gov/?id=EJ949445>

From the abstract: This article examines the language proficiency and achievement outcomes of Latino students enrolled in a dual language programme who varied by language proficiency (Native English speakers, Current English Language Learners--ELLs, Fluent English Proficient/Previous ELLs). Most previous research has not disaggregated Latino students, especially ELLs. The purpose of this research is to examine the achievement and language proficiency of 732 Grade 4 to Grade 8 Latino students enrolled in a dual language programme who differed by language proficiency. Results show that these Latino student groups achieve at higher levels than their peers in English mainstream. Findings also indicated that the three groups vary in parent education, language proficiency in Spanish, and achievement as measured in Spanish and English. Further, Fluent English Proficient/Previous ELLs are the most Spanish proficient and bilingual, achieve at higher levels in English and Spanish, and close the achievement gap with native English speakers in English mainstream programmes.

Ong, S. L. (2013). Usefulness of dual-language science test for bilingual learners. *Studies in Educational Evaluation*, 39(2), 82–89. NO URL

From the abstract: This study examines the usefulness of the science test presented in a dual-language format in two separate science test booklets, one comprising English-only test items and the other, dual-language test items. The participants were 1720 eight-grade students from 26 secondary schools. Most of the students viewed the dual-language test positively as they felt it enhanced their understanding of the test items. However, only two items were found to

function significantly different in the dual-language format. Students' performance for the two versions of the test was comparable. The results showed that the extra language version did not provide greater accessibility and comprehensibility of the test to the students. The findings may prove valuable to decision-making regarding language accommodation policies for testing in content areas.

Sanchez, S. V., Rodriguez, B. J., Soto-Huerta, M. E., Villarreal, F. C., Guerra, N. S., & Flores, B. B. (2013). A case for multidimensional bilingual assessment. *Language Assessment Quarterly*, 10(2), 160–177. <https://eric.ed.gov/?id=EJ1003318>

From the abstract: Current assessment practices in the United States are not able to accurately capture the total linguistic, cognitive, and achievement abilities of bilingual learners. There are psychometric complexities involved when assessing and interpreting test results of bilingual students, which impact the validity of this practice. Further, the compromise associated with measuring bilingual students in only one of their two languages has been found to produce a distorted picture, one that has contributed to the overrepresentation of bilingual students in special education programs. This study presents case data using a multidimensional bilingual assessment approach that provides evidence against a single language assessment approach. The results reveal the complexity associated with measuring bilingual students' skills as well as the quandary that is introduced. The case data demonstrate the importance of a multidimensional bilingual assessment that begins with determining a student's cognitive and academic language proficiency. The case data also demonstrate how the reliability and validity of other assessments may be impacted by the unique language development trajectories exhibited by bilingual learners. The study concludes with the recommendation to provide a multidimensional bilingual assessment, which will maximize the reliability and validity of results and provide teachers with the benefit of information in both languages that can then be used to facilitate instructional supports as well as link to meaningful instruction and interventions.

Spencer-Iiams, J. (2013). *Passage reading fluency in Spanish and English: The relation to state assessment outcomes in English for students in a dual-language context* (Order No. 3589569). Available from Education Collection. (1430909864). <https://eric.ed.gov/?id=ED559080>

From the abstract: “The United States is experiencing an increase in young students developing literacy in English and Spanish. Schools providing dual-language English/Spanish instruction need technically adequate tools to assess reading skills in the

languages of instruction, and interpretation of results needs to acknowledge the complexity of cross-linguistic learning. Although passage reading fluency in English strongly predicts overall reading proficiency in English in the primary grades and there is some indication that passage reading fluency in Spanish provides equivalent information regarding Spanish reading skills, rarely have the two been examined simultaneously and within a dual-language instructional context. The current study examined predictive and concurrent validity of passage reading fluency in English and Spanish within third grade within a dual-language instructional environment. Using a state assessment of reading as the criterion measure, a correlational design was used to investigate the relation between passage reading fluency in English and Spanish and performance on the statewide assessment of reading in English. Findings indicate that within a dual-language context, passage reading fluency in English is the stronger predictor of performance on the state assessment in English, regardless of the student's home language. Spanish reading fluency is also strongly related to English reading fluency but did not explain additional variance in predicting performance on the statewide large-scale assessment of reading in English beyond what English fluency explained. Results are consistent with the idea that same language assessments are more predictive of reading performance than cross-language assessments are, but the benefits of formative assessment in the language of instruction remain."

Methods

Keywords and Search Strings

The following keywords and search strings were used to search the reference databases and other sources:

- "language" AND "assessment"
- "language of assessment"
- "language of assessment" NOT "dissertations and theses"
- "English language learners" AND "content assessment"
- "dual language" AND "content assessment"
- "bilingual students" AND "assessment"
- "native" AND "language" AND "assessment"
- "native" AND "language" AND "assessment" AND "pacific"
- "dual language" AND "assessment" AND "pacific"
- "dual language" AND "assessment" NOT "dissertations and theses"
- "dual language" AND "test" NOT "dissertations and theses"
- "English as a foreign language" AND "assessment" NOT "dissertations and theses"
- "English as a foreign language" AND "test" NOT "dissertations and theses"

Databases and Resources

ERIC, EBSCO Host, ProQuest Education Journals, Google/Google Scholar

Reference Search and Selection Criteria

REL Pacific searched ERIC and other academic journal databases for studies that were published in English-language peer-reviewed research journals within the last 20 years. REL Pacific prioritized documents that are accessible online, although not all sources may be publicly available, and prioritized references that provide practical information based on peer-reviewed research for the teachers and leaders who inquired about multi-grade classrooms for this Ask A REL.¹ Resources included in this document—including URLs, descriptions, and content—were last accessed in June

¹ This memorandum is one in a series of quick-turnaround responses to specific questions posed by educational stakeholders in the Pacific Region (American Samoa, the Commonwealth of the Northern Mariana Islands, the Federated States of Micronesia, Guam, Hawai'i, the Republic of the Marshall Islands, and the Republic of Palau), which is served by the Regional Educational Laboratory (REL Pacific) at McREL International. This memorandum was prepared by REL Pacific under a contract with the U.S. Department of Education's Institute of Education Sciences (IES), Contract ED-IES-17-C-0010, administered by McREL International. Its content does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.