

Prepared for LESC August 2012

Presented By

Dean Michael A. Morehead and Dr. Ron Dziwenka

New Mexico State University

## Ways in Which Testing or Assessment Misinforms, Misrepresents &/or is Misused Concerning VAMs & ELLS

Tests are a tool, not a goal. We should use them as needed, not let them use us. The more we rely on **high-stakes standardized tests**, the more we **destroy students' creativity, ingenuity, and willingness to think differently**, and the more we **demoralize teachers**.<sup>i</sup> **Diane Ravitch**<sup>ii</sup> critiques today's reformers' most popular ideas for restructuring schools, **including privatization, standardized testing, and punitive accountability**, and shows why **the business model is not an appropriate way to improve schools**. **Linda Darling-Hammond** holds that these and other home, school, and student factors influence student learning gains, and these matter more than the individual teacher in explaining changes in scores.<sup>iii</sup> She<sup>iv</sup> asserts that **value-added evaluation hurts teaching**, and cites key findings to support this view in studies by the NRC & NAS,<sup>v</sup> RAND,<sup>vi</sup> and ETS.<sup>vii</sup> There is an overemphasis on the value of test scores over more influential variables to student learning. Most importantly, these test scores largely reflect whom a teacher teaches, not how well they teach. First, **test-score gains reflect much more than an individual teacher's effort**, including students' health, home life, and school attendance, and schools' class sizes, curriculum materials, and administrative supports, as well as the influence of other teachers, tutors, and specialists. Second, teachers' **ratings are highly unstable**: They differ substantially across classes, tests, and years. Third, **teachers who rate highest on the low-level multiple-choice tests** currently in use are **often not those who raise scores on assessments of more-challenging learning**.

### A. Value-Added Measures (VAMs)

According to Koertz, overreliance on standardized tests runs the **risk of excessive or inappropriate teaching to the test**.<sup>viii</sup> In Chetty, Friedman & Rockoff's<sup>ix</sup> controversial recent study on VAMs, they assert that **evaluating teachers using test scores could encourage, or pressure** people to engage in, **counterproductive behaviors**. Due to NCLB waivers, administrators doing the evaluating are **likely to be looking for secondary indicators that teachers are**, in fact, **preparing their students for tests**. An important limitation of C, F & R's

analysis is that teachers they studied were not incentivized based on student test scores.<sup>x</sup> They refute the argument that since states are only using test scores for PART of a teacher's evaluation (NY state requires that **40% of an eval be derived from test scores**), teachers will NOT feel much pressure to teach to the tests. However, in a release from the State Education Department, "Teachers rated ineffective on student performance based on objective assessments must be rated ineffective overall." Thus, Ravitch tweeted: "Teacher in NY agreement **rated "ineffective" on 40% (test scores) will be rated ineffective, period. So 40%=100%.**" Basing 40% of a teacher's evaluation on test scores will indeed promote teaching to the test. Even when models try to control for prior achievement and student demographic variables, **teachers** are advantaged or **disadvantaged based on the students they teach**. Teachers with large numbers of new ELLs and others with special needs show lower gains than the same teachers when they are teaching other students.<sup>xi</sup> The **pattern of lower scores in classes with large numbers of ELLs is well known**. Penalizing a teacher for taking on the toughest assignment does not make sense.

Standardized test scores **offer an incomplete picture of learning**.<sup>xii</sup> A test is **merely a small sample of behavior** that we use to estimate mastery of a **much larger "domain" of (student) achievement**; it is not the domain. Overreliance on testing lends itself to **validating entirely arbitrary performance targets**.<sup>xiii</sup> Thus, but it is not possible to use this model to determine whether a given level of performance is desirable. Also, because each student's expected score is largely derived from the student's actual scores in previous years, a school with high levels of student turnover may have difficulty in collecting sufficient VAM data. They also caution against putting too much emphasis on what might be a **weak correlation between the test score impact on longer-run outcomes**. Test scores alone offer **only a snapshot**, and should mainly be used as mechanisms for support, offering teachers feedback pathways to improvement.<sup>xiv</sup>

The 2012 Primary Sources study<sup>xv</sup> shows that most **teachers** support using multiple measures of student achievement, and say that **standardized tests do not accurately reflect their students' growth**.<sup>xvi</sup> And, according to an NWEA study,<sup>xvii</sup> **formative and interim assessments** are perceived by parents, educators and administrators as **more valuable than summative assessments**. The reformers championing the VAM model employ the same **logical fallacy, circular reasoning**, that doomed NCLB. They have defined great teaching as that which results in the most gains on end of year tests, and then spent millions of dollars identifying indicators of teaching that will yield the best scores. The most deceptive strategy is how they then **try to pretend that these indicators are "multiple measures" of good teaching. In fact, these are simply indicators of teaching practices associated with higher test scores**. A Tennessee SCORE study,<sup>xviii</sup> found that the state was evaluating teachers (50% of teachers' evaluations are based on student testing data) using the data of students many didn't even teach, raising credibility issues concerning the (excessive) use of student test scores for teacher evaluation.

VAM's **"Big Data" can be flawed, or at best unvalidated, data**. Estimates of growth in individual classrooms as well as VAMs of teacher effectiveness **in a single year are generally very imprecise and highly unstable**. Schochet & Chiang<sup>xix</sup> show that the **error rate for comparing teacher performance with one year of data is likely to be 35%** (65% accuracy, in the world of U.S. education, is a D). The result is that in any given year, many teachers will be misclassified. A second problem is that the **statistical models employed are complex**, and the **education field has not yet agreed which methods are best**.<sup>xx</sup> Test questions, or (portions of)

the test itself, can be of **questionable quality, validity, or reliability**, as can what they (are supposed to) measure. Recent examples include the ELA exam “**Pineapple-gate**”<sup>xxi</sup> issue of Pearson, the largest test publisher in the world, and of Florida's 2012 Science FCAT Test.<sup>xxii</sup> These examples highlight the fact that in today's data-driven world of high-stakes standardized testing, the scores students achieve (which which are used to evaluate in part teachers and principals) even in a flawed test could be used to end a teacher's career. According to Todd Farley,<sup>xxiii</sup> we know that testing data can be manipulated to tell any story. Any school administration—by **making test questions easier** or **lowering cut scores**—can **portray improvement** in its classrooms even when such improvement doesn't really exist.

This is an example of the misuse of Item Analysis (IA). In Item Analysis, test items are continuously refined in order to come to results in which 50% are “perfect” and 50% are not – **goal is to discriminate among respondents / responses.... NOT to find out what respondents know**. In addition, Popham discusses in detail the issue of **item facility (IF)**. Test items with **poor IFs** tend to be **dropped since they adversely impact score distributions**. Market forces impel test developers to create tests with high reliability indices, for which well-spread score distributions are requisite.<sup>xxiv</sup>

**Test-based rewards** and sanctions are **not** having the **desired effect on student and school outcomes**, and are slowing the nation's progress in **closing the achievement gap**.<sup>xxv</sup> According to 600,000 member New York State United Teachers (NYSUT) President Richard C. Iannuzzi,<sup>xxvi</sup> more than 2,000 delegates to its Representative Assembly in Buffalo adopted a resolution urging the State Education Department to reduce the focus on questionable standardized tests in favor of other measures of student learning that are more "accurate, fair and appropriate." Study after study suggest **economic inequity is built into, and worsened by, school systems**, so why do self-styled education "reformers" keep ignoring class issues and instead focus on standardized testing? For example, Rothwell's study<sup>xxvii</sup> demonstrates that location greatly affects test scores.

**Assessment-driven accountability requirements, and structuring the curriculum around standardized testing**, thus relying excessively on student testing, **drains the love of learning from students and the love of teaching from teachers**. Kevin Wilner, director of the National Education Policy Center, laments that the recent 2012 MetLife Survey<sup>xxviii</sup> shows that teacher job satisfaction has dropped alarmingly, 15 percentage points, just over the past two years. Teachers have been watching sadly as the sort of **engaging learning** that attracted them to the profession is **increasingly squeezed out**. **Testing places a floor on whatever skills are** (supposedly) **measured by the required test**, and this might deter high-quality applicants from teaching in public schools. Moreover, test (certification) requirements may disqualify some applicants that schools would otherwise want to hire.<sup>xxix</sup>

.....

## B. English Language Learners (ELLs) & Testing

According to the 2000 US Census, almost one out of every five Americans speaks a language besides English at home (p. 19) and the US Dept of Ed (2006) predicts that by 2025, ELLs will constitute 25% of all school-aged students (p. 21). **ELLs are nearly 11 percent of the K-12 population**, and about **80 percent** of these students **speak Spanish**, with the rest speaking a wide variety of other languages. The rest speak one of more than 400 languages, and over half of all ELLs were born in the U.S.<sup>xxx</sup> Menken states ‘**testing and accountability** under the law ultimately **reflect a “language-as-problem”** or “deficit model” **orientation in recent U.S. language policy**, where **language has become a liability for ELLs**’ (p. 160), and ‘**when test scores are attached to high-stakes decisions** like high school graduation, they **can limit the future opportunities of ELLs**’ (p. 184).

Every year, some 5 million public school students who are still mastering English take assessments. Unfortunately, the **results of these tests are far from valid** because many of these students are not sufficiently proficient in English to demonstrate their knowledge and abilities on assessments designed for native English-speakers.<sup>xxxii</sup> ELLs placed in national and state assessment and accountability systems<sup>xxxii</sup> can be placed at a **disadvantage** because **assessment outcomes may not be valid** due to the impact of their **limited English proficiency on content knowledge** performance, **ELLs may** not have received the same curriculum as non-ELLs and are **tested on content** for which they have **not received instruction**; and **assessment tools in large-scale assessments** usually **constructed for native speakers** of English **may be biased** toward these students. **Current measures of English proficiency** may be **classifying** students as English proficient based on **oral proficiency**, which may **not guarantee readiness** to succeed in English-only classrooms (Collier 1995, Francis et al. 2006a, Genesee et al., 2006, Hakuta et al. 2000, Moore and Zainuddin 2003, Oakely et al. 1998).

The **vast majority** of high-stakes **tests are written and administered only in English**. This often leaves ELLs at a disadvantage and raises questions as to how the test results should be interpreted.<sup>xxxiii</sup> When ELLs take **standardized tests**, the **results tend to reflect their English language proficiency** and **may not accurately assess their content knowledge or skills**,<sup>xxxiv</sup> therefore **weakening the test’s validity** for them. However, **tests in languages other than English** are **rarely provided**.<sup>xxxv</sup> Robert Linquanti<sup>xxxvi</sup> points out that the interrelationship of the two goals of proficiency in academic content and the English language poses significant challenges to current assessment and accountability policies. The long-term ELLs studied by Menken & Kleyn<sup>xxxvii</sup> rarely had the opportunity to hone their native language skills, even though research shows that **literacy skills** students learn **in their native languages transfer to English** (Cummins, 2000).<sup>xxxviii</sup> As well, ELL students with **formal schooling in their first language** tend to **acquire English proficiency faster** than their peers without it (Collier 1995, Garcia-Vazques et al 1997, Genesee et al. 2006).<sup>xxxix</sup>

According to David Plank,<sup>xl</sup> **the language-learner subgroup is wrongly constituted**, merely for accountability purposes. The most linguistically and academically accomplished ELL students exit the English-language-learner category over time, as they become fluent in English. Those not making sufficient progress remain, where they are joined by newly entering ELLs who

are by definition at lower levels of language proficiency. State-level **assessment results** typically **ignore these revolving-door practices**, which wrongly **stigmatize** the language-learner subgroup,<sup>xii</sup> **demoralize** students and teachers, and **prevent accurate reporting of long-term outcomes**, including **graduation rates and college access and success** data. Tremendous inconsistencies in the identification and classification of ELs affect the validity, accuracy, and comparability of outcome data.<sup>xiii</sup> **Federal law requires** that ELLs be provided with **accommodations but**, decisions on the **number and type of accommodations** to be used with ELLs are **left to each state**.<sup>xiii</sup>

Numerous studies have documented the fact **that language & performance on achievement tests are confounded for ELLs** more than for most students.<sup>xiv</sup> In **mathematics and science**, **test items** may have complex **linguistic structures unrelated** to the focal construct that unnecessarily add to cognitive load and slow the reader down.<sup>xv</sup> Studies have found **linguistic modifications** of test questions with excessive language demands to be **effective**, by demonstrating that the **unnecessary linguistic complexity of content-based assessments** (e.g., mathematics and science) is a **likely source of measurement error differentially affecting** the reliability and validity of assessments for the **ELL subgroup**.<sup>xvi</sup> Since **almost all assessments measure language proficiency to some degree**, ELLs may receive lower scores on content area assessments administered in English than they would if they took the same tests in a language in which they were proficient.<sup>xvii</sup> **The longer** that the **students are in the school system, the harder** it can be for them **to show proficiency in English**, because **academic standards become more rigorous** as they move up through the grades. **In the long run**, they are lacking academic literacy. They don't always understand what the teacher says because the teacher is speaking the **academic language that they lack**.<sup>xviii</sup> Results of analyses of extant data from (NAEP) suggested that **ELLs had difficulty with the linguistically complex test items**. Studies also found that ELLs exhibited a substantially higher number of omitted or not-reached test items.<sup>xix</sup> The results of analyses of existing data from several locations nationwide show a substantial gap in reliability (internal consistency) and validity (concurrent validity) between ELLs and non-ELLs on test items with a substantial language demand.<sup>1</sup> In language, science, and social science, the gap on alpha between English-only and ELLs was large.<sup>li</sup>

According to Pandya's work,<sup>lii</sup> **the current overemphasis on high stakes testing and accountability fails ELLs by imposing & reinforcing a common set of beliefs about poor people (ELLs)**, instead of instilling a sense of the complexity of ELL and the variety of strengths & weaknesses of the range of ELLs in today's classrooms. Benesch<sup>liii</sup> provides a postmodern critique of the construct of Generation 1.5. Specifically she notes the presence of three partialities in the discourse surrounding Generation 1.5: (1) the partiality of demographics, (2) the partiality of language, and (3) the partiality of academics. She argues that these three partialities continue to **perpetuate the monocultural/monolingual ideology** of the United States and serve **to present the identities of members Generation 1.5 as in-between or deficient rather than as fluid and multiple**. As well, ELLs are included in high-stakes tests in which the **cultural familiarity and knowledge is assumed**. Test items may contain **references to ideas or events that are unfamiliar to ELLs** because they have not been exposed to similar concepts in their native culture and have not lived in the United States for a long period of time.<sup>liv</sup>

According to Rivera and others,<sup>lv</sup> the **accommodations most frequently used for ELLs are timing/scheduling and setting**, and not presentation (repeating questions or translating) or response (allowing ELLs to respond in their native language). While allowing an ELL more time to complete a test or administering the test in a smaller group in familiar surroundings may be helpful in some contexts, such **accommodations do not ensure that learners' linguistic needs are being accounted for**. "You are taking a test not crafted for English language learners and trying to retrofit it through accommodations. It is a real Band-Aid." The **inequitable distribution of instructional resources to appropriately support ELLs' learning** and the substantial need for better preparation, coaching, and ongoing professional development of all teachers of ELLs, make it all the more important to develop ELL-relevant formative assessment processes and practices that can provide feedback and guide next steps in teaching and learning for linguistic and academic growth.<sup>lvi</sup> Researchers noted that the **number of studies that met their criteria for inclusion was small**.<sup>lvii</sup>

## BIBLIOGRAPHY

- Abedi, Jamal, "High-stakes Tests, English Language Learners, and Linguistic Modification," *Sunshine State TESOL Journal*, Vol 6, No. 1, Spring 2007.
- Abedi, J. (2006). Language Issues in Item-Development. In Downing, S. M. and Haladyna, T. M. *Handbook of Test Development* (Ed.). New Jersey: Lawrence Erlbaum Associates, Publishers.
- Abedi, J., Leon, S., Mirocha, J. (2003). *Impact of students' language background on content-based data: Analyses of extant data* (CSE Tech. Rep. No. 603). UCLA: Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Student Testing.
- Abedi, J., Courtney, M., & Leon, S. (2003). *Effectiveness and validity of accommodations for English language learners in large-scale assessments* (CSE Tech. Rep. No. 608). 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(3), 219-234.
- 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.
- Angrist, Joshua D. & Guryan, Jonathan, "Does teacher testing raise teacher quality? Evidence from state certification requirements," 2007.
- Benesch, S. (2008). *Generation 1.5 and its Discourses of Partiality: A Critical Analysis*. *Journal of Language, Identity, and Education*, 7(3-4), 294-311.

Callahan, R. (2006). The Intersection of Accountability and Language: Can Reading Intervention Replace English Language Development? *Bilingual Research Journal*, 30(1), 1-21.

Capps, R., Fix, M., Murray, J., Ost, J., Passel, J. S., & Herwanto, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act*. Washington, DC: Urban Institute.

Chapman, Ben & Monahan, Rachael, "Talking pineapple question on state exam stumps ... everyone!" /NEW YORK DAILY NEWS, April 19, 2012.

Chapman, Matt. "For every child, multiple measures: what parents & educators want from K-12 assessments," NWEA, 2012.

Chetty, Raj, Friedman, John N., & Rockoff, Jonah E., "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood," NBER, Cambridge, Mass., Dec 2011.

Chevalier, J. (2004). Heritage Language Literacy: Theory and Practice. *The Heritage Language Journal*, 2(1), 1-19.

Cody, Anthony, "Pineapplegate Raises Fresh Questions about the Obama Administration's Expansion of Testing," Education Week, April 23, 2012.

Cody, Anthony, "The Pineapple Story Tests Us: Have Test Publishers become Unquestionable Authorities?" Education Week, April 20, 2012.

Coltrane, Bronwyn, "English Language Learners and High-Stakes Tests: An Overview of the Issues," Center for Applied Linguistics, Nov. 2002.

Darling-Hammond, Linda, "Value-Added Evaluation Hurts Teaching," Education Week (March 5, 2012).

Darling-Hammond, et.al "Getting Teacher Evaluation Right – Background Paper for policy makers," AERA & NAE, Sept, 2011.

ETS. "Using Student Progress to Evaluate Teachers - A Primer on Value-Added Models," 2005.

Farley, Todd, *Making the Grades: My Misadventures in the Standardized Testing Industry*. Berrett-Koehler Publishers, October 1, 2009.

Florida Dept of Education, "Florida Writing Tests Set Off Alarms; State Board Sets New Cut Score," (State EdWatch Blog) May 16, 2012.

Gándara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English learners in California Schools: Unequal Resources; Unequal Outcomes. *Educational Policy Analysis Archives*. <http://.epaa.asu.edu/epaa/v11n36/>

“Guidelines for the Assessment of ELLs,” Educational Testing Service, 2009.

Goldenberg, C. (2008, Summer). Teaching English Language Learners: What the Research Does-and Does not-Say. *American Educator*, 8-44.

Hawkins, Beth. “Designer of Value-Added Tests a Skeptic About Current Test Mania,” *Education Week*, June 19, 2012.

Hawkins, Beth. “Student-testing pioneer Angermeyr is skeptical about high-stakes trends,” *Minnesota Post*, June 15, 2012.

Jacob, B and S. D. Levitt, S.D., "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," NBER Working Paper No. [9413](#), January 2003, and *Quarterly Journal of Economics*, 117 (August 2003), pp. 843-77.

Katz, S. R. (1996). Where the Streets Cross the Classroom: A Study of Latino Students' Perspectives on Cultural Identity in City Schools and Neighborhood Gangs. *Bilingual Research Journal*, 20(3 & 4), 603-631.

Kieffer, M., Lesaux, N., Rivera, M. & Francis, D. (2009). Accommodations for English-language learners taking large scale assessments: A meta-analyses on effectiveness and validity. *Review of Educational Research*, 79(3), 1168-1201.

Koertz, Daniel. *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press, First Edition, 2008.

Kopriva R. (Ed.). (2008). *Improving testing for English Language learners*. New York: Routledge Press.

Letter Report to the U.S. Department of Education on the Race (NRC & NAS - Oct 2009).

*Linquanti, Robert*, “Strengthening Assessment for English Learner Success: How Can the Promise of the Common Core Standards and Innovative Assessment Systems Be Realized?” In *The Road Ahead for State Assessments*, PACE, 2011.

Menken, Kate and Kleyn, Tatyana, “The Difficult Road for Long-Term English Learners,” in *Supporting English Language Learners*, Volume 66, Number 7, April 2009.

Menken, K. *English Learners Left Behind: Standardized Testing as Language Policy*. Clevedon, UK: Multilingual Matters, 2008.

Menken, K. (2000). *What are the critical issues in wide-scale assessment of English language learners?* (Issue Brief No. 6). Washington, DC: National Clearinghouse for Bilingual Education. Retrieved July 7, 2012, from <http://www.ncela.gwu.edu/ncbepubs/issuebriefs/ib6.htm>

MetLife Survey of the American Teacher: Teachers, Parents & the Economy, Metlife, Inc., March, 2012.

McCaffrey, Lockwood, Koretz & Hamilton, “Evaluating Value-Added Models for Teacher Accountability” (RAND 2003).

The National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs (NCELA), 2007, [www.ncela.gwu.edu/expert/fastfaq/4.html](http://www.ncela.gwu.edu/expert/fastfaq/4.html)

NET3 (National Evaluation of Title III) (2010, February). Preliminary data from the NET3 provided to the U.S. Department of Education by the American Institutes for Research.

“NYSUT demands end to 'broken' testing system,” NYSUT Media Relations site – accessed April 27, 2012.

Plank, David N., Assessing English-Language Learners: One Size Does Not Fit All, Education Week, Aug. 31, 2011.

Popham, W. James, *The Truth about Testing: An Educator's Call to Action*. Alexandria, VA: Association for Supervision and Curriculum Development , 2001.

“Primary Sources 2012: America's teachers on the teaching profession,” Scholastic, and The Bill and Melinda Gates Foundation, 2012.

Ravitch, Diane. “The Problem Is Bigger Than a Pineapple,” *Education Week*, April 24, 2012.

Ravitch, D. *The Death and Life of the Great American School System: How Testing and Choice are Undermining Education*, ReadHowYouWant, 2012.

Rivera, Francis, D., Lesaux, M., Kieffer, N., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for instruction and academic interventions*. Portsmouth, NH: RMC Research Corporation, Center on Instruction. Available: [www.centeroninstruction.org/files/ELL1-Interventions.pdf](http://www.centeroninstruction.org/files/ELL1-Interventions.pdf)

Rivera, C., Stansfield, C., Scialdone, L., & Sharkey, M. (2000). *An analysis of state policies for the inclusion and accommodation of ELLs in state assessment programs during 1998-1999* (Executive Summary). Washington, DC: The George Washington University, Center for Equity and Excellence in Education.

Rothwell, Jonathan. “Housing Costs, Zoning, and Access to High-Scoring Schools,” Metropolitan Policy Program at Brookings, 2012.

Schochet, Peter Z & Chiang, Hanley S. "Error Rates in Measuring Teacher & School Performance Based on Student Test Score Gains." *Mathematica Policy Research*, July 2010.

Shaftel, J., Belton-Kocher, E., Glasnapp, D., & Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of English language learners and students with disabilities. *Educational Assessment*, 11(2), 105-126.

Solano-Flores, G. (2006). Language, dialect, and register: Sociolinguistics and the estimation of measurement error in the testing of English language learners. *Teachers College Record* (108)11, pp. 2354-2379.

Strauss, Valerie. "Why faith in standardized testing industry is misplaced," *Washington Post*, April 11, 2011.

Taylor, J., Stecher, R., O'Day, J., Naftel, S. & LeFloch, K.C. (2010). *State and Local Implementation of the No Child Left Behind Act, Volume IX – Accountability under NCLB: Final Report*. Washington, DC: U.S. Department of Education.

"Teacher Survey Shows Morale Is at a Low Point," *The New York Times*, Mar. 8.

Ucelli-Kashyap, Marla. "Fix the Fixation on Testing," in *Education Week* (May 22, 2012).

Willner, Shafer, Rivera, L. C., & Acosta, B. (2008). Descriptive study of state assessment policies for accommodating English language learners. Prepared for the LEP Partnership, U.S. Department of Education. Arlington, VA: The George Washington University Center for Equity and Excellence in Education.

The Working Group on ELL Policy (2009). *The American Recovery and Reinvestment Act: Recommendations for Addressing the Needs of English Language Learners*. Palo Alto, Calif.: Author. Retrieved on July 9, 2012, from <http://www.stanford.edu/~hakuta/ARRA/>

Young, J. W., & King, T. C. (2008). *Testing accommodations for English language learners: A review of state and district policies*: New York: The College Board. Report # 2008-6.

Zacher Pandya, Jessica, *Overtested: How High-Stakes Accountability Fails English Language Learners*, Teachers College Press, Columbia University, 2011.

---

<sup>i</sup> Diane Ravitch, "The Problem Is Bigger Than a Pineapple," *Education Week*, April 24, 2012.

<sup>ii</sup> Ravitch, D. *The Death and Life of the Great American School System: How Testing and Choice are Undermining Education*, ReadHowYouWant, 2012. She suggests that there is "something fundamentally antidemocratic about relinquishing control of the public educational policy to private foundations run by society's wealthiest people."

<sup>iii</sup> Darling-Hammond, et.al "Getting Teacher Evaluation Right – Background Paper for policy makers," AERA & NAE, Sept, 2011.

<sup>iv</sup> Darling-Hammond, Linda, "Value-Added Evaluation Hurts Teaching," *Education Week* (March 5, 2012).

<sup>v</sup> Letter Report to the U.S. Department of Education on the Race (NRC & NAS - Oct 2009).

<sup>vi</sup> McCaffrey, Lockwood, Koretz & Hamilton, "Evaluating Value-Added Models for Teacher Accountability" (RAND 2003).

<sup>vii</sup> "Using Student Progress to Evaluate Teachers - A Primer on Value-Added Models," ETS, 2005.

<sup>viii</sup> Koertz, Daniel. *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press, First Edition, 2008. If you focus too much on the tested sample, at the expense of the broader domain it represents, you get inflated scores. Scores no longer represent real achievement. It's better to focus on the knowledge and skills the test is intended to represent, not the details of particular test items. No one measure can evaluate (teacher or student) performance. Daniel Koertz is a Harvard professor who specializes in educational measurement.

<sup>ix</sup> Raj Chetty and John N. Friedman, economists at Harvard (along with Jonah E. Rockoff of Columbia) co-wrote the much-discussed recent study "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood."

<sup>x</sup> In C, F & R's sample, the top 2% of teachers ranked have patterns of test score gains that are consistent with test manipulation, based on the proxy developed by Jacob and Levitt. They caution that the link between teacher value-added and students' life outcomes might change once testing becomes high stakes for teachers themselves. See *B. Jacob and S. D. Levitt, "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," NBER Working Paper No. 9413, January 2003, and Quarterly Journal of Economics, 117 (August 2003), pp. 843-77.*

<sup>xi</sup> Darling-Hammond, et.al "Getting Teacher Evaluation Right – Background Paper for policy makers," AERA & NAE, Sept, 2011.

<sup>xii</sup> Schooling has goals other than achievement. Tests provide limited, specialized information about student performance. Even a very good test measures only a modest proportion of what we value. Koertz, Daniel. *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press, First Edition, 2008.

<sup>xiii</sup> **Jim Angermeyr has recently spoken out on the recent trend of overusing VAMs. He is one of the architects of the value-added assessment, is not so thrilled. Two decades ago, Jim was one of a small group of educators and psychologists who came up with a new kind of standardized test, the growth model or value-added assessment. He worked with the Northwest Evaluation Association to develop tests, and more recently as director of research and evaluation with the Bloomington Public Schools. Where the distortion comes in is that you can only test a limited amount of the domain. When you can summarize a whole bunch of complicated things in a single number, that has a lot of power and it's hard to ignore, especially when it tells a story that you want to promote. And that's where it gets really twisted. Testing professionals know that you're just sampling the domain and you don't try to make inferences further than that. But nonprofessionals do that all the time. I would do away with standards. They assume all kids are the same and they all make progress the same way and move in lockstep. And that's just not accurate. Standards distort individual differences among kids. And that's bad. The amount of money we spend every year giving reading and math tests to students to give them a high school diploma has done absolutely nothing to improve the graduation rate; it's done nothing to improve the quality of the graduates. See Beth Hawkins, "Student-testing pioneer Angermeyr is skeptical about high-stakes trends," Minnesota Post, June 15, 2012; and "Designer of Value-Added Tests a Skeptic About Current Test Mania," Education Week, June 19, 2012.**

<sup>xiv</sup> Raj Chetty and John N. Friedman, economists at Harvard (along with Jonah E. Rockoff of Columbia) co-wrote the much-discussed recent study "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood."

<sup>xv</sup> As a survey of more than 10,000 public school teachers, it found that only 28% of educators see state-required standardized tests as an essential or very important gauge of student achievement. In addition, only 26% of teachers say standardized tests are an accurate reflection of what students know. Only 45% of teachers say their students take such tests seriously or perform to the best of their ability on them. Fully 85% of teachers agree that student growth over the school year should be a part of a teachers' evaluation! However only 36% agree that we should rely on standardized tests for this data. "Primary Sources 2012: America's teachers on the teaching profession," Scholastic, and The Bill and Melinda Gates Foundation, 2012.

<sup>xvi</sup> We delude ourselves into thinking we have measured learning because we uncritically accept the premise that 'learning is measurable'. Answering a question with a correct choice does not mean one has correct understanding. Not only can one guess, but also he can have wrong reasons for the correct answer.

<sup>xvii</sup> Key findings from the study include: Child-centered teaching and learning is a top priority for parents and educators; parents, teachers and district administrators think it's important to measure student performance in a full range of subjects; and many parents, teachers and administrators question the money, time and stress spent on assessment. See Chapman, Matt. "For every child, multiple measures: what parents & educators want from K-12 assessments," NWEA, 2012.

<sup>xviii</sup> Only about one-third teach subjects where value-added testing data is collected. The SCORE (State Collaborative on Reforming Education) report recommends that teachers in subjects or grades without specific testing data be allowed to reduce that component to 25% of their evaluation. "The state knew all along that the lack of test data for the majority of teachers in the state was a huge weakness in the system, but they stubbornly moved forward with the evaluations this past year anyway," TEA lobbyist Jerry Winters said in an email. "The use of school-wide test data—evaluating teachers using the data of students they may not even teach—is a blatant mistake which raises major credibility issues." "Supporting Effective Teaching in Tennessee - listening & gathering feedback on Tennessee's teacher evaluations," SCORE, 2012.

<sup>xix</sup> Schochet, Peter Z & Chiang, Hanley S. "Error Rates in Measuring Teacher & School Performance Based on Student Test Score Gains." Mathematica Policy Research, July 2010.

<sup>xx</sup> We can't be confident about the share of that growth that is properly attributed to the effects of the teacher. And value-added methods do nothing whatever to address the core problems of poorly designed test-based accountability: inappropriate test prep and score inflation. Koertz, Daniel. *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press, First Edition, 2008.

In state after state, 20 to 50% of teacher-effectiveness ratings are now determined by such data. However, evaluation schemes based on student-achievement data produce inconsistent results from year to year. Teachers' ratings differ substantially from *class to class* and from *year to year*, as well as from one *test* to the next. See Linda Darling-Hammond, et al "Getting Teacher Evaluation Right – Background Paper for policy makers," AERA & NAE, Sept, 2011.

<sup>xxi</sup> The story is an absurd tale of a talking pineapple, who challenges a hare to a race. When the pineapple fails to move and the rabbit wins, the animals dine on the pineapple. Students were asked two perplexing questions: why did the animals eat the talking fruit, and which animal was wisest? Teachers, principals and parents said they weren't sure what the answers were. See Diane Ravitch, "The Problem Is Bigger Than a Pineapple," Education Week, April 24, 2012; Anthony Cody, "The Pineapple Story Tests Us: Have Test Publishers become Unquestionable Authorities?" Education Week, April 20, 2012; BEN CHAPMAN AND RACHEL MONAHAN, "Talking pineapple question on state exam stumps ... everyone!" /NEW YORK DAILY NEWS, April 19, 2012; and Anthony Cody, "Pineapplegate Raises Fresh Questions about the Obama Administration's Expansion of Testing," Education Week, April 23, 2012.

<sup>xxii</sup> The percentage of 4th grade students with passing scores on the writing portion of the Florida 2012 FCAT plunged from 81% last year to 27% this year, making it look as if most students went from good to horrible writers in one year. The board realized, after the fact, that student writing wasn't really any worse, but the new test-scoring guide was too harsh and penalized students for minor mistakes. The sudden drop in scores called the state's entire testing system into question. See "2012 FCAT writing changes & preliminary results," Florida Dept of Education, May 2012. See also "Florida Writing Tests Set Off Alarms; State Board Sets New Cut Score," (State EdWatch Blog) May 16, 2012.

<sup>xxiii</sup> Todd Farley is the author of "Making the Grades: My Misadventures in the Standardized Testing Industry." Farley worked for 15 years in the K-12 testing business for many of the biggest players (Pearson Education, Educational Testing Service, American Institutes of Research, etc.) on many of the biggest tests (National Assessment of Educational Progress, California High School Exit Exam, Florida Comprehensive Assessment, Virginia Standards of Learning, etc.). According to Farley, we know that the **testing companies fudge numbers all the time**, whether **reliability numbers** (to show the industry is doing a more "standardized" job than it really is); **validity numbers** (to show the industry is doing a more accurate job than it really is); or **score distribution numbers** (when test scoring companies work to ensure student results match the predictions of their own psychometricians). Farley writes, there seems to be a major disconnect between the profit motive of the testing industry's major players (Pearson Education, McGraw-Hill, Riverside Publishing, ETS K-12, DRC...) and any altruistic goals for American education.

In 2000, a scoring error by NCS-Pearson (now Pearson Educational Measurement) led to 8,000 Minnesota students being told they failed a state math test when they did not, in fact, fail it (some of those students weren't able to graduate from high school on time). In 2004, ETS erroneously informed over 4,000 teachers they had failed a PRAXIS exam that they had actually passed, leading to lost jobs and lawsuits aplenty. In 2006 Pearson again erred, giving lower scores than were deserved to more than 4,000 students taking the SAT, with the company making the excuse (apparently with a straight face) that their blunder resulted due to "abnormally high moisture content" in that year's score sheets. See Strauss, Valerie. "Why faith in standardized testing industry is misplaced," Washington Post, April 11, 2011.

<sup>xxiv</sup> Popham, W. James, *The Truth about Testing: An Educator's Call to Action*. Alexandria, VA: Association for Supervision and Curriculum Development, 2001, p. 48.

<sup>xxv</sup> Ucelli-Kashyap, Marla. "Fix the Fixation on Testing," in Education Week (May 22, 2012). Ucelli-Kashyap is Assistant to the President for Educational Issues, American Federation of Teachers.

<sup>xxvi</sup> Findings include teachers reported students becoming restless over the excessive length of the tests, too many administrators, concerned about the ramifications of lower-than-expected scores, are stressing "test prep" at the expense of real learning, the state's over-reliance on testing is undermining parental support, and concerns that factors beyond the control of public schools, such as poverty, hunger, student attendance, parental involvement and the lack of community resources for schools, were being sidelined. See "NYSUT demands end to 'broken' testing system," NYSUT Media Relations site – accessed April 27, 2012.

<sup>xxvii</sup> Rothwell, Jonathan. "Housing Costs, Zoning, and Access to High-Scoring Schools," Metropolitan Policy Program at Brookings, 2012.

<sup>xxviii</sup> MetLife Survey of the American Teacher: Teachers, Parents & the Economy, Metlife, Inc., March, 2012. See also "Teacher Survey Shows Morale Is at a Low Point," The New York Times, Mar. 8.

<sup>xxix</sup> Angrist, Joshua D. & Guryan, Jonathan, "Does teacher testing raise teacher quality? Evidence from state certification requirements," 2007.

<sup>xxx</sup> Capps, R., Fix, M., Murray, J., Ost, J., Passel, J. S., & Herwanto, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act*. Washington, DC: Urban Institute; and The National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs (NCELA), 2007, [www.ncele.gwu.edu/expert/fastfaq/4.html](http://www.ncele.gwu.edu/expert/fastfaq/4.html)

<sup>xxx</sup> Abedi, Jamal, "High-stakes Tests, English Language Learners, and Linguistic Modification," *Sunshine State TESOL Journal*, Vol 6, No. 1, Spring 2007; Abedi, J. (2006). Language Issues in Item-Development. In Downing, S. M. and Haladyna, T. M. *Handbook of Test Development* (Ed.). New Jersey: Lawrence Erlbaum Associates, Publishers.

<sup>xxxii</sup> Abedi, Jamal, "High-stakes Tests, English Language Learners, and Linguistic Modification," *Sunshine State TESOL Journal*, Vol 6, No. 1, Spring 2007.

<sup>xxxiii</sup> Historically, ELLs were not included in high-stakes standardized tests (Lara & August, 1996). This practice resulted in a lack of accountability for the academic progress of ELLs, with ELLs not being held to the same high academic standards as their peers. Consequently, **ELLs did not benefit from the educational reforms** that followed the implementation of high-stakes assessments (August & Hakuta, 1997). Bronwyn Coltrane, "English Language Learners and High-Stakes Tests: An Overview of the Issues," Center for Applied Linguistics, Nov. 2002.

<sup>xxxiv</sup> Menken, K. (2000). *What are the critical issues in wide-scale assessment of English language learners?* (Issue Brief No. 6). Washington, DC: National Clearinghouse for Bilingual Education. Retrieved July 7, 2012, from <http://www.ncele.gwu.edu/ncbepubs/issuebriefs/ib6.htm>

<sup>xxxv</sup> In fact, **testing accommodations** that involve **translation** of a test into a student's native language are **frequently prohibited by states**. Rivera, C., Stansfield, C., Scialdone, L., & Sharkey, M. (2000). *An analysis of state policies for the inclusion and accommodation of ELLs in state assessment programs during 1998-1999* (Executive Summary). Washington, DC: The George Washington University, Center for Equity and Excellence in Education.

<sup>xxxvi</sup> First, **academic assessments that fail to take account of ELs' English-language proficiency level will likely inadequately measure their content area knowledge** and skills, both individually and as a subgroup. Students with limited proficiency in English often underperform on assessments of academic content, reflecting not a lack of knowledge but a lack of fluency, which may unfairly depress their scores. Second, current accountability policies **distort the overall EL cohort's academic performance** and obscure long-term outcomes by removing those who achieve English proficiency from the EL subgroup. **As a result, reporting of subgroup academic performance** is limited to those EL students who are **by definition the lowest performing**. See Robert Linqunti, "Strengthening Assessment for English Learner Success: How Can the Promise of the Common Core Standards and Innovative Assessment Systems Be Realized?" In *The Road Ahead for State Assessments*, PACE, 2011, p. 13-26.

<sup>xxxvii</sup> This is despite the fact that research has shown that **native language literacy ability** is one of the **greatest predictors of academic performance in school for English language learners** (Ford, 2005; Thomas & Collier, 1997). Research is also conclusive that teaching students to read in their native language promotes higher levels of reading achievement in English (Goldenberg, 2008; Baker, 2006; Krashen & McField, 2005; Thomas & Collier, 1997). Kate Menken and Tatyana Kleyn, "The Difficult Road for Long-Term English Learners," in *Supporting English Language Learners*, Volume 66, Number 7, April 2009.

<sup>xxxviii</sup> See also Callahan, R. (2006). The Intersection of Accountability and Language: Can Reading Intervention Replace English Language Development? *Bilingual Research Journal*, 30(1), 1-21; Goldenberg, C. (2008, Summer). Teaching English Language Learners: What the Research Does-and Does not-Say. *American Educator*, 8-44.

<sup>xxxix</sup> On the other hand, ELL students who **attend high-poverty schools** tend to **acquire English at slower rate** than other ELLs. (Abedi and Dietel 2004, Hakuta et al. 2000, Jepsen and de Alth 2005).

Educational Testing Service (ETS) suggests the **main factors influencing the assessment of ELLs** are: language (different linguistic backgrounds and varying levels of proficiency in English as well as in their native languages), educational background (varying degrees of formal schooling in their languages and in English, and varying degrees of exposure to standardized testing), and cultural (varying degrees of acculturation to U.S. mainstream). See "Guidelines for the Assessment of ELLs," Educational Testing Service, 2009.

<sup>xl</sup> David N. Plank, *Assessing English-Language Learners: One Size Does Not Fit All*, Education Week, Aug. 31, 2011. David N. Plank is the executive director of Policy Analysis for California Education, or PACE, an independent, nonpartisan research center based at the University of California, Berkeley; the University of Southern California; and Stanford University. Before joining PACE in January 2007, he was a professor at Michigan State University, where he founded and directed the Education Policy Center. Mr. Plank is the author or editor of six books, including the American Educational Research Association's *Handbook of Education Policy Research* (2009).

<sup>xli</sup> What's more, states have little incentive to track the long-term performance after the two years of required monitoring. See also Katz, S. R. (1996). Where the Streets Cross the Classroom: A Study of Latino Students' Perspectives on Cultural Identity in City Schools and Neighborhood Gangs. *Bilingual Research Journal*, 20(3 & 4), 603-631.

<sup>xlii</sup> NET3 (National Evaluation of Title III) (2010, February). Preliminary data from the NET3 provided to the U.S. Department of Education by the American Institutes for Research.; and The Working Group on ELL Policy (2009). *The American Recovery and Reinvestment Act: Recommendations for Addressing the Needs of English Language Learners*. Palo Alto, Calif.: Author. Retrieved on July 9, 2012, from <http://www.stanford.edu/~hakuta/ARRA/>.

<sup>xliii</sup> Kieffer, M., Lesaux, N., Rivera, M. & Francis, D. (2009). Accommodations for English-language learners taking large scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research*, 79(3), 1168-1201; Shafer Willner, L., Rivera, C., & Acosta, B. (2008). Descriptive study of state assessment policies for accommodating English language learners. Prepared for the LEP Partnership, U.S. Department of Education. Arlington, VA: The George Washington University Center for Equity and Excellence in Education; Solano-Flores, G. (2006). Language, dialect, and register: Sociolinguistics and the estimation of measurement error in the testing of English language learners. *Teachers College Record* (108)11, pp. 2354-2379; and Young, J. W., & King, T. C. (2008). *Testing accommodations for English language learners: A review of state and district policies*: New York: The College Board. Report # 2008-6.

<sup>xliiv</sup> 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; and Abedi, J., Courtney, M., & Leon, S. (2003). *Effectiveness and validity of accommodations for English language learners in large-scale assessments* (CSE Tech. Rep. No. 608). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

<sup>xlv</sup> 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; Abedi, J. (2006). Language Issues in Item-Development. In Downing, S. M. and Haladyna, T. M. (Ed.), *Handbook of Test Development* (pp. 377-398). New Jersey: Lawrence Erlbaum Associates, Publishers; Kopriva R. (Ed.). (2008). *Improving testing for English Language learners*. New York: Routledge Press; Shaftel, J., Belton-Kocher, E., Glasnapp, D., & Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of English language learners and students with disabilities. *Educational Assessment*, 11(2), 105-126; and Solano-Flores, G. (2006). Language, dialect, and register: Sociolinguistics and the estimation of measurement error in the testing of English language learners. *Teachers College Record* (108)11, pp. 2354-2379.

<sup>xlvi</sup> Abedi, J. & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, 14(3), 219-234.

<sup>xlvii</sup> **For example, an ELL who has the mathematical skills needed to solve a word problem may fail to understand the task because of limited English proficiency. In this case, the assessment is testing not only mathematical ability, but also English proficiency. If the construct of interest is mathematical skill exclusive of language skills, then it may be systematically inaccurate to base inferences about the academic content knowledge or skills of this student and other ELLs on the scores of tests delivered in English. This distinction can be complicated if the construct of interest is not merely mathematical skill, but rather the ability to do mathematics within an English-medium classroom. See "Guidelines for the Assessment of ELLs," Educational Testing Service, 2009.**

<sup>xlviii</sup> See also Chevalier, J. (2004). Heritage Language Literacy: Theory and Practice. *The Heritage Language Journal*, 2(1), 1-19.

<sup>xlix</sup> 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.

<sup>l</sup> Abedi, J. (2006). Language Issues in Item-Development. In Downing, S. M. and Haladyna, T. M. *Handbook of Test Development* (Ed.). New Jersey: Lawrence Erlbaum Associates, Publishers; and Abedi, J., Leon, S., Mirocha, J. (2003). *Impact of students' language background on content-based data: Analyses of extant data* (CSE Tech. Rep. No. 603). UCLA: Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Student Testing.

<sup>li</sup> Not only ELLs, but native speakers of English at the lower tail of the academic performance distribution may also benefit from assessment with simple linguistic structure. Abedi, Courtney, Leon, Kao, and Azzam, in press.

<sup>lii</sup> Jessica Zacher Pandya, *Overtested: How High-Stakes Accountability Fails English Language Learners*, Teachers College Press, Columbia University, 2011.

<sup>liii</sup> Benesch, S. (2008). *Generation 1.5 and its Discourses of Partiality: A Critical Analysis*. *Journal of Language, Identity, and Education*, 7(3-4), 294-311.

<sup>liv</sup> **ELLs may not be familiar with the kind of language that is used in tests, including many predictable patterns and phrases. It may also be beneficial to teach test-taking skills (e.g., how to approach a multiple-choice question, how to locate the main idea in a reading passage) to help prepare ELLs for specific types of test items they may encounter. Bronwyn Coltrane, "English Language Learners and High-Stakes Tests: An Overview of the Issues," Center for Applied Linguistics, Nov. 2002.**

---

<sup>lv</sup> Rivera, C., Stansfield, C., Scialdone, L., & Sharkey, M. (2000). *An analysis of state policies for the inclusion and accommodation of ELLs in state assessment programs during 1998-1999* (Executive Summary). Washington, DC: The George Washington University, Center for Equity and Excellence in Education.

<sup>lvi</sup> Gándara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English learners in California Schools: Unequal Resources; Unequal Outcomes. *Educational Policy Analysis Archives*. <http://epaa.asu.edu/epaa/v11n36/>; and Taylor, J., Stecher, R., O'Day, J., Naftel, S. & LeFloch, K.C. (2010). *State and Local Implementation of the No Child Left Behind Act, Volume IX – Accountability under NCLB: Final Report*. Washington, DC: U.S. Department of Education.

<sup>lvii</sup> Francis, D., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for instruction and academic interventions*. Portsmouth, NH: RMC Research Corporation, Center on Instruction. Available: [www.centeroninstruction.org/files/ELL1-Interventions.pdf](http://www.centeroninstruction.org/files/ELL1-Interventions.pdf)