



Title of Project:

Assessing English Language Learners' Content Knowledge
in the Mainstream Assessing ELL content knowledge

Researcher:

Beth Clark-Gareca
New York University
bc210@nyu.edu



Beth Clark-Gareca

Research Supervisor:

Dr. Lorena Llosa

Background to the Study

In the current educational climate in the U.S., elementary students are taking more tests than ever before. English language learners (ELLs) represent one of the fastest growing demographics in public schools and have been greatly affected by new assessment policies because the language of the tests presents them with significant linguistic challenges (Abedi, Lord, Hofstetter, & Baker, 2000; Menken, 2008; Rivera & Collum, 2006a). Test accommodations are one way to increase the likelihood that test scores reflect ELLs' content mastery rather than their level of language proficiency, yet apart from the work of Lee and her colleagues in science content (Lee, 2004; Lee, Maerten-Rivera, Buxton, Penfield, & Secada, 2009) relatively few ELL accommodations studies have been conducted in classroom contexts. In fact, to date there has been no empirical examination of exactly how, or indeed *if*, routine content tests are typically modified at all (Cizek, 2007).

The purpose of this dissertation study was to make a contribution to the understanding of how ELLs' math and science content knowledge is routinely assessed in mainstream elementary classrooms. The research questions explored the assessment practices and beliefs of elementary teachers, as well as ELLs' perceptions of the classroom assessments in which they participated. A mixed method design was utilized incorporating multiple data sources, including a wide-scale teacher survey (n=213), classroom observations, teacher interviews, and student interviews with math and science teachers (n=10) and fourth grade ELLs (n=50).

Findings

Given the broad scale of this study, the main findings are summarized here in brief. Results from the survey suggested that the majority of classroom content teachers (65%) were not certain of the proficiency levels of the ELLs in their classrooms, nor were they certain which of their students were identified as requiring English as a Second Language (ESL) services. ELLs who were dually identified as having Special Education and ESL needs were found to consistently receive more accommodations than any other group. Twenty-one classroom assessment practices were observed and reported in use in the ten focal classrooms. Practices that were universally implemented were: additional time, teacher assistance, and reading items aloud. Practices which



were reported to be commonly implemented were: shortened tests, time management by the teacher, collaborative test-taking, and paraphrasing of test items. Practices that were infrequently implemented in both classroom and survey report were native language accommodations such as oral translation, English-English dictionary use, bilingual dictionary use, and bilingual tests. For accommodations to be effective for ELLs, they need to be linguistically appropriate (Acosta, Rivera, & Shafer Willner, 2008); however, appropriate native language accommodations were found not to be implemented routinely in this study. Four accommodated grading practices: grade modification, explicit item correction, retesting/retaking assessments, and test exemption, were also identified in the focal classrooms as being typical of teachers' practice. Grade elevation was universally reported in all ten classrooms as a normal grading practice. To illustrate this point, one teacher explained his grading practices for ELLs thus, "*Sometimes I just simply go by how my heart feels. I honestly do that.*"

The beliefs underlying teachers' decision-making processes were found to be a determining factor in ELL assessment, and teacher accounts of ELLs' performance were often found to be in direct contrast with ELLs' perceptions of the same assessment events. All ten focal teachers expressed a belief that some aspect of ELLs' home lives, i.e., a lack of parental commitment or a general lack of English at home, had a negative impact on students' academic success. In contrast to this common teacher belief, a large number of children reported speaking some English at home (40%), living in stable family environments with their mother, father, or both parents (92%), and working hard to meet their parents' expectations for them (74%). One child, Tomy, explained his Dad's reaction if he brought home a bad grade, "*No, he wouldn't be very happy, but he'll still love me.*" These findings echo long-standing work related to lowered teacher expectations for at-risk populations, such as ELLs and students of color (Brown, Benkovitz, Muttillo, Urban, 2005; Delpit, 1995; Ladson-Billings, 1997; Valenzuela, 1999). They also reinforce assertions related to a deficit perspective commonly associated with ELLs (Pettit, 2011). Misalignment of teacher and student accounts of assessment events was noted throughout the study and suggested that misunderstanding may be an integral part to classroom assessment of ELLs.

Teacher beliefs related to ELLs' language proficiency tended to fall into two categories: students who could participate in regular classroom instruction and assessment and students who could not. Teachers identified oral proficiency in English as a major indicator of students' ability to function in fourth grade content classrooms, a finding which supports previous research that students at the lowest levels of English proficiency have the greatest need for accommodations (Acosta, Rivera, Shafer-Willner, 2008). In fact, eight of ten teachers reported encountering assessment difficulties with students who could not yet communicate orally at a high enough language proficiency level. In terms of language, many students expressed a limited understanding of their own language processes; that is, despite being bilingual, the children lacked metacognitive skills for talking about their language acquisition processes. Thirty students (60%) expressed a belief that their developing English language did not affect their ability to do well on content tests in spite of the fact that many were still at beginner and intermediate levels of language proficiency. When asked what it would be like to take a test in their native languages, thirty students (60%) said that they thought it would be difficult, not because they



didn't know the content of the tests, but because they didn't know how to read in their native languages. Magdalena expressed the opinions of many students as to why translated tests would be difficult, *"It would be like kinda harder to read because I'm like here, and now sometimes I forget how to read in Spanish."*

Although much literature has focused on children's aversion to tests and the negative effect that assessment can have on learning environments (Jones, et al., 1999; Mulvenon, Stegman, & Ritter, 2005), a majority of ELLs in this study reported that they liked taking tests. Seventy-eight percent of ELL participants reported that they liked math tests and 64% reported to like science tests. Of the fifty ELL participants, 22 (44%) reported that tests presented them with opportunities to learn, recognizing assessments as formative tools designed for their learning. When students were asked about the accommodations in place for them, forty-six (92%) claimed to like them, giving reasons such as accommodations helped them to learn or earn better scores. Only four children (8%) spoke of any concerns related to the stigma of accommodations implementation; two of those children spoke to the stigma of being exempted from regular assessment that their non-ELL peers were required to take. When asked about fairness of accommodations, several ELLs said that they appreciated being given extra time, though they thought it was unfair that their English-speaking classmates were not afforded the same opportunity. Many students also reported that trying was an important factor in doing well on tests, supporting previous findings in work conducted relating to effort and motivation with native English speaking students (Brookhart, 2007; Brookhart & DeVoge, 1999; Brookhart, Walsh & Zientarski, 2006).

A majority of students spoke positively about teachers' modified grading practices because they typically resulted in higher test scores; nonetheless, 70% of students did not understand some aspect of how their work was evaluated by their teachers. Because accommodated scoring practices were found to be frequent among all teachers there was a lack of alignment between classroom grades and state standardized test scores, which caused a number of ELLs in the study genuine concern. Competing rating systems used to grade students' work were also reported to be the cause of significant confusion, and students in the study reported that B could mean *Basic* on a high-stakes rating scale, *Good* on a standard A through F rating scale, or *Bien* on rating scales from their countries of origin. Additionally, the use of percentages to rate fourth graders presented conceptual difficulty since percentages as a mathematical concept was taught as part of the fourth grade math curriculum.

Implications of the Study

In sum, numerous implications of this study were characterized along three strands: 1) implications for professional development, 2) implications for assessment, and 3) implications for policy. These strands have great consequences for the academic success of ELLs and serve as critical points of consideration in going forward.

To answer Cizek's (2007) call, increased teacher professional development in the area of assessment is necessary for teachers so they can better evaluate ELLs at varying proficiency levels on classroom assessments. A stronger understanding of the basic principles of assessment,



and the need for systematic implementation and withdrawal of accommodations for ELLs could help content teachers develop greater expertise when assessing ELLs' content knowledge. Further teacher training in multicultural education and second language acquisition was requested by a majority of participant teachers to better inform their practice, and could function to dispel some of the misinformation found to be held by teachers about ELLs and their communities.

In reference to assessment theory, the variable implementation of accommodations and the prevalence of accommodated grading practices for ELLs suggest that “a hodgepodge of factors” come into play when teachers assess student performance (McMillan, 2004, p. 203). These haphazard systems can cause content test scores, and subsequent grades and school decisions to be an inaccurate account of ELL student content knowledge. Students need to be accountable to their own learning through measurable outcomes, and consistent and transparent systems of assessment are essential to this process. Without such systems in place, grades become an imprecise and ultimately meaningless measure of student performance. To that end, school based grading systems and test use should be reconsidered according to constructs of assessment theory in order to facilitate a more informed discussion of the place of classroom tests and the role of classroom accommodations as a whole.

The results of this study also have many implications for policy. High-stakes testing procedures, which rely on student familiarity with accommodations from classroom tests, are likely misguided, because many test accommodations offered on high-stakes tests were not found to be implemented in the classroom to a great extent. In accordance with other work done relating to marginalization of ELL populations in school (Menken, 2006; Pettit, 2011), de facto school policies supporting the placement of ELLs into the lowest content tracks should be revisited to ensure that ELLs receive the most appropriate education that they are entitled to by law. In the schools that had multiple content sections, ELLs were found in great numbers in the lowest tracks where accommodations were administered to all content learners on a whole-class basis. Whole class accommodations policies represent a conflation of ELLs with children in Special Education or with other learning needs, and cannot optimally serve students in either population. Results from this study suggest that teachers need more guidance to understand and attend to the unique linguistic needs that ELLs have when engaging in content assessment.



References

- Abedi, J. (2000). Confounding of students' performance and their language background variables. Report: (ERIC Document Reproduction Service No. ED449250).
- Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, 33(1), 4-14.
- Abedi, J. (2006). Psychometric issues in the ELL assessment and special education eligibility. *Teachers College Record*, 108(11), 2282-2303.
- Abedi, J., Courtney, M., & Leon, S. (2003). Research-supported accommodation for English language learners in NAEP. (CSE Tech Rep. No., 586). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Abedi, J. & Dietel, R. (2004). Challenges in the No Child Left Behind Act for English language learners. *Phi Delta Kappan*, 85(10), 782.
- Abedi, J., Hofstetter, C., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy based empirical research. *Review of Educational Research*, 74 (1), 1-28.
- Abedi, J. & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, 14(3), 219-134.
- Abedi, J., Lord, C., & Hofstetter, C. (1998). Impact of selected background variables on students' NAEP math performance (CSE Tech. Rep. No. 478). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- 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.
- Abrams, L. (2007). Implications of high-stakes testing for the use of formative classroom assessment. In J. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 79-98). New York, NY: Teachers College Press.
- Acosta, B., Rivera, C., & Shafer-Willner, L. (2008). *Best practices in state assessment policies for accommodating English language learners: A Delphi study*. Retrieved January 2, 2013, from ceee.gwu.edu/AA/BestPractices.pdf.
- Albus, D., Thurlow, M., Liu, K., Bielinski, J. (2005). Reading test performance of English language learners using an English dictionary. *The Journal of Educational Research*, 98 (4), 245-254.



- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1985). *Standards for educational and psychological testing*, Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999). *Standards for educational and psychological testing*, Washington, DC: American Educational Research Association.
- Anderson, N. E., Jenkins, F. F., & Miller, K. E. (1996). NAEP inclusion criteria and testing accommodations: Findings from the NAEP 1995 field test in mathematics. Washington, DC: Educational Testing Service.
- Appleton, K. (2002). Science activities that work: Perceptions of primary school teachers. *Research in Science Education*, 32(3), 393-410.
- Assessment Reform Group (1999). *Assessment for learning, Beyond the black box*. Cambridge: Cambridge University Press.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Ballard, R. & Elwark, A. (2003). Learning disability and professional licensing examinations: what accommodations are reasonable under the ADA? *The Journal of Psychiatry and Law*, 31, 43-66.
- Barr, S. (2012). Core strategies to support English language learners. *The Educational Forum*, 76, 105-117.
- Batt, E. (2008). Teachers' perceptions of ELL education: Potential solutions to overcome the greatest challenges. *Multicultural Education*: 15(3), 39-43.
- Beswick, J., Willms, J., & Sloat, E. (2005). A comparative study of teacher ratings of emergent literacy skills and student performance on a standardized measure. *Education*, 126(1), 116-137.
- Bloom, B., & Krathwohl, D. (1956). *Taxonomy of educational objectives: the classification of educational goals*. New York, Longman.



- Bogdan, R., & Biklen, S. (1998). *Qualitative research for education: an introduction to theory and methods*. Boston, MA: Allyn and Bacon.
- Breen, M., Hird, B., Milton, M., Oliver, R., & Thwaite, A. (2001). Making sense of language teaching: teachers' principles and classroom practices. *Applied Linguistics*, 22(4), 470-501.
- Brinckerhoff, L., & Banerjee, M. (2007). Misconceptions regarding accommodations on high-stakes tests: Recommendations for preparing disability documentation for test takers with learning disabilities. *Learning Disabilities Research and Practice*, 32(4), 246-255.
- Brindley, G. (1998). Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing*, 15(1), 45-85.
- Brislin, R. (1970). Back-translation for cross-cultural research. *Journal of Cross Cultural Psychology*, 1(3), 185-216.
- Brookhart, S. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education*, 10, 161-180.
- Brookhart, S. (2007). Expanding views about formative classroom assessment: A review of the literature. In J. McMillan (Ed.), *Formative classroom assessment: Theory into practice*. (pp. 43-62). New York, NY: Teachers College Press.
- Brookhart, S., Andolina, M., Zuza, M., & Furman, R. (2004). Minute math: An action research study of pupil self-assessment. *Educational Studies in Mathematics*, 57, 213-227.
- Brookhart, S., & DeVoge, J. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education*, 12(4), 409-25.
- Brookhart, S., Walsh, J., & Zientarski, W. (2006). The dynamics of motivation and effort for classroom assessments in middle school science and social studies. *Applied Measurement in Education*, 19(2), 151-184.
- Brown, K., Benkovitz, J. Mutillo, A. J., & Urban, T. (2005). Leading schools of excellence and equity: Documenting effective strategies in closing achievement gaps. *Teachers College Record*, 113(1), 57-96.
- Butler, F. & Stevens, R. (1997). Accommodation strategies for English language learners on large scale assessment: Student characteristics and other considerations. (CSE Tech. Rep. No. 448). Los Angeles: CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).



- Butler, F. & Stevens, R. (2001). Standardized assessment of the content knowledge of English language learners K-12: Current trends and old dilemmas. *Language Testing*, 18 (4), 409-427.
- Byrnes, D., Kiger, G. & Manning, M. (1997). Teachers' attitudes about language diversity. *Journal of Applied Social Psychology*, 26, 455-467.
- Carlo, M. S., August, D., Mclaughlin, B., Snow, C. E., Dressler, C., & Lippman, D. N. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classroom. *Reading Research Quarterly*, 39(2), 188-215.
- Castellon-Wellington, M. (2000). The impact of preference for accommodations: The performance of English language learners on large scale academic achievement tests. CSE Technical Report 524. Los Angeles: University of California, Center for the Study of Evaluation.
- Cheng, L., Klinger, D., & Zheng, Y. (2007). The challenges of the Ontario Secondary School Literacy Test for second language students. *Language Testing*, 24(185), 185-208.
- Cho, S., & Reich, G. (2008). New immigrants, new challenges: High school social studies teachers and English language learner instruction. *The Social Studies Journal*, 99(6), 235-242.
- Cizek, G. (2007). Formative classroom and large-scale assessment: Implications for future research and development. In J. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 99-115). New York, NY: Teachers College Press.
- Clair, N. (1995) Mainstream classroom teachers and ESL students. *TESOL Quarterly*, 29(1) 189-196.
- Cloud, N. (2003). Special education needs of second language students. In J. Richards (Ed.), *Educating second language children: The whole child, the whole curriculum, the whole community* (pp. 243-277). Cambridge: Cambridge University Press.
- Common Core State Standards Initiative (2013). *Implementing the common core state standards*. Retrieved January 16, 2013, from <http://www.corestandards.org>.
- Crawford, J. (2001). *At war with diversity: US language policy in an age of anxiety*. Clevedon, UK: Multilingual Matters.
- Creswell, J. (2007). *Qualitative inquiry and research design*. (2nd ed). Thousand Oaks, CA: Sage Publications.



- Crosby, E., & French, J. (2002). Psychometric data for teacher judgments regarding the learning behaviors of primary grade children. *Psychology in the Schools*, 39(3), 235-244.
- Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Clevedon, England: Multilingual Matters.
- Curtin, E. (2005). Instructional styles used by regular classroom teachers while teaching recently mainstreamed ESL students: Six urban middle school teachers in Texas share their experiences and perceptions. *Multicultural Education*, Sum, 36-42.
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 305-34.
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43(3), 393-415.
- deJong, E., & Harper, C. (2005). Preparing mainstream teachers for English language learners: Is being a good teacher good enough? *Teacher Education Quarterly*, 32(2), 101- 124.
- Delpit, L. (1995). *Other people's children: Cultural conflict in the classroom*. New York, NY: Norton.
- DeStefano, L., Shriner, J., & Lloyd, C. (2001). Teacher decision making in participation of students with disabilities in large-scale assessment. *Exceptional Children*, 68(1), 7-23.
- DeVincenzi, F. (1995). Language tests and ESL teaching - Examining standardized test content: Some advice for teachers. *TESOL Quarterly*, 29(1), 180-184.
- Diaz-Rico, L. & Weed, K. (2002). *The cross-cultural language, and academic development handbook*. (2nd ed.) Boston, MA: Allyn and Bacon.
- Diaz-Rico, L., & Weed, K. (2009). *The cross-cultural language, and academic development handbook*. (4th ed). Boston, MA: Allyn and Bacon.
- Dixon, H., Hawe, E., & Parr, J. (2011) Enacting assessment for learning: The beliefs practice nexus. *Assessment in Education: Principles, Policy & Practice*, 18(4), 365-379.
- Doe, C., Cheng, L., Fox, J., Klinger, D., & Zheng, Y. (2011). What has experience got to do with it? An exploration of L1 and L2 test-takers' perceptions of test performance and alignment to classroom literacy activities. *Canadian Journal of Education*, 34(3), 68-85.
- Donovan, M. & Cross, C. (Eds.). (2002). *Minority students in special and gifted education*. Washington DC: National Academy Press.



- Duncan, T., Parent, L., Chen, W., Ferrara, S., Johnson, E., Oppler, S., & Shieh, Y. (2005). Study of a dual language test booklet in eighth-grade mathematics. *Applied Measurement in Education* 18(2), 129-161.
- Eckert, T., Dunn, E., Coddling, R., Begeny, J., & Kleinmann, A. (2006). Assessment of mathematics and reading performance: An examination of the correspondence between direct assessment of student performance and teacher report. *Psychology in the Schools*, 43(3), 247-65.
- Eder, D., & Fingerson, L. (2003). Interviewing children and adolescents. In J. Holstein & F. Gubrium (Eds.), *Inside Interviewing: new lenses, new concerns* (pp. 33-55). London: Sage Publications.
- Escamilla, K., Chavez, L., & Vigil, P. (2005) Rethinking the “Gap”: High-stakes testing and Spanish speaking students in Colorado. *Journal of Teacher Education*, 56(2), 132-144.
- 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.
- Freeman, Y., & Freeman, D. (1998). *ESL/EFL teaching: Principles for success*. Portsmouth, NH: Heinemann.
- Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Seabury Press.
- Fuchs, L., & Fuchs, D. (2001). Helping teachers formulate sound test accommodation decisions for students with learning disabilities. *Learning Disabilities: Research and Practice* 16(3), 174-81.
- Fuchs, L., Fuchs, D., & Capizzi, A. (2005). Identifying appropriate test accommodations for students with learning disabilities. *Focus on Exceptional Children*: 37(6), 1-9.
- Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlett, C. L., Binkley, E., & Crouch, R. (2000a). Using objective data sources to enhance teacher judgments about test accommodations. *Exceptional Children*, 67, 67-81.
- Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlett, C. L., & Karns, K. (2000b). Supplementing teacher judgments of mathematics test accommodations with objective data sources. *School Psychology Review*, 29, 65-85.
- Garcia, G. (1991). Factors influencing the English reading test performance of Spanish speaking Hispanic children. *Reading Research Quarterly*, 14(4), 371-92.



- Garcia, T., del Rio Paraent, L., Chen, L., Ferrara, S., Garavaglia, D., Johnson, E., Linag, J., Oppler, S., Searcy, C., Shieh, Y., & Ye, Y. (2000). *Study of a dual language test booklet in 8th grade mathematics: Final report*. Washington DC: American Institutes for Research.
- Gersten, R. (1999). Lost opportunities: Challenges confronting four teachers of English language learners. *The Elementary School Journal*, 100(1), 37-56.
- Gijssel, M., Bosman, A., & Verhoeven, L. (2006). Kindergarten risk factors, cognitive factors, and teacher judgments as predictors of early reading in Dutch. *Journal of Learning Disabilities*, 39(6), 558-571.
- Giraud, G., & Impara, J. (2005). Making the cut: The cut score setting process in a public school district. *Applied Measurement in Education*, 18(3), 289-312.
- Gottlieb, M. (2006). *Assessing English language learners: Bridges from language proficiency to academic achievement*. Thousand Oaks, CA: Corwin Press.
- Hafner, A. (2000). *Evaluating the impact of test accommodations on test scores of LEP students and non-LEP students*. Dover, DE: Delaware Department of Education.
- Hafner, A. (2001). *Evaluating the impact of test accommodations on test scores of LEP students and non-LEP students*. Paper presented at annual meeting of the American educational research association, Seattle, WA.
- Harklau, L. (1994). ESL versus mainstream classes: Contrasting L2 learning environments. *TESOL Quarterly*, 28(2), 241-272.
- Harper, C., & deJong, E. (2004). Misconceptions about teaching English language learners. *Journal of Adolescent and Adult Literacy*, 48(2) 152-162.
- Hecht, S., & Greenfield, D. (2002). Explaining the predictive accuracy of teacher judgments of their students' reading achievement: The role of gender, classroom behavior, and emergent literacy skills in a longitudinal sample of children exposed to poverty. *Reading and Writing: An Interdisciplinary Journal*, 15(7-8), 789-809.
- Helwig, R., & Tindal, G. (2003). An experimental analysis of accommodation decisions on large scale mathematics tests. *Exceptional Children*, 69(2), 211-225.
- Herrera, S., Murry, K., & Cabral, R. (2007). *Assessment accommodations for classroom teachers of culturally and linguistically diverse students*. Boston, MA: Pearson.
- Hofstetter, C. (2003). Contextual and mathematics accommodation test effects for English language learners. *Applied Measurement in Education*, 16(2), 159-188.



- Hollenbeck, K., Tindal, G., & Almond, P. (1998). Teachers' knowledge of accommodations as a validity issue in high stakes testing. *Journal of Special Education*, 32(3), 175-83.
- Jones, B., & Egley, R. (2007). Learning to take tests of learning for understanding? Teachers' beliefs about test-based accountability. *The Educational Forum*, 71, 232-48.
- Jones, M. G., Jones, B., Hardin, B., Chapman, L., Yarbrough, T. & Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81(3), 199-203.
- Jones, M., Jones, B., & Hargrove, T. (2003). *The unintended consequences of high-stakes testing*. Lanham, MD: Rowman & Littlefield.
- Karabenick, S., & Noda, P. (2004). Professional development implications of teachers' belief and attitudes toward English language learners. *Bilingual Research Journal*, 28(1), 55-75.
- Ketterlin-Geller, L., Alonzo, J., Braun-Monegan, J., & Tindal, (2007). Recommendations for accommodations: Implications of (in)consistency. *Remedial and Special Education*, 28(4), 194-206.
- Ketterlin-Geller, L. (2005). Knowing what all students know: Procedures for developing universal design for assessment. *Journal of Technology, Learning, and Assessment*, 4(2). Retrieved April 12, 2013, from <http://files.eric.ed.gov/fulltext/EJ848519.pdf>
- 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.
- King, M., & Youngs, P. (2003). Classroom teachers' views on inclusion. *Research Institute on Secondary Education Reform for Youth with Disabilities (RISER) Brief*. Wisconsin Center for Education Research, No. 7, 1-14.
- Kiplinger, V.L., Haug, C.A., & Abedi, J. (2000). *A math assessment should test math, not reading: One state's approach to the problem*. Paper presented at the 30th annual national conference on large-scale assessment, Snowbird, UT.
- Koenig, J. A., & Bachman, L. F. (Eds.). (2004). *Keeping score for all: The effects of inclusion and accommodation policies on large-scale educational assessments*. Washington, DC: National Academies Press.
- Kopriva, R. (2000). *Ensuring accuracy in testing for English language learners*. Council of chief state school officers, SCASS, LEP Consortium, 1-115.



- Kopriva, R., Emick, J., & Hipolito-Delgado, C. (2007). Do proper accommodation assignments make a difference? Examining the impact of improved decision making on scores for English language learners. *Educational Measurement Issues and Practice*, 26(3), 11-20.
- Kunnan, A. (1999). Recent developments in language testing. *Annual Review of Applied Linguistics*, 19, 235-53.
- 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(1), 55-75.
- Ladson-Billings, G. (1997). It doesn't add up: African American students' mathematics achievement. *Journal for Research in Mathematics Education*, 28(6), 697-708.
- Lager, C. (2006). Types of mathematics-language reading interactions that unnecessarily hinder algebra learning and assessment. *Reading Psychology*, 27(2), 165-204.
- Larson, C., & Ovando, C. (2001). *The color of bureaucracy: The politics of equity in multicultural school communities*. Belmont, CA: Wadsworth.
- Lee, J. S., & Oxelson, E. (2006). "It's not my job" K-12 teachers' attitudes toward students' heritage language maintenance. *Bilingual Research Journal*, 30(2), 453-477.
- Lee, O. (2003). Equity for linguistically and culturally diverse students in science education: A research agenda. *Teachers College Record*, 105(3), 465-89.
- Lee, O. (2004). Teacher change in beliefs and practices in science and literary instruction with English language learners. *Journal of Research in Science Teaching*, 41(1), 65-93.
- Lee, O., & Fradd, S. (1998). Science for all, including students from non-English-language backgrounds. *Educational Researcher*, 27(4), 12-21.
- Lee, O., Hart, J., Cuevas, P., & Enders, C. (2004). Professional development in inquiry based science for elementary school teachers of diverse student groups. *Journal of Research in Science Teaching*. 41(10), 1021-1043.
- Lee, O., Maerten-Rivera, J., Buxton, C., Penfield, R., & Secada, W. (2009). Urban elementary teachers' perspectives on teaching science to English language learners. *Journal of Science Teacher Education*, 20(3), 263-86.
- Lee, O., Penfield, R., & Buxton, C. (2011). Relationship between "form" and "content" in science writing among English language learners. *Teachers College Record*, 113(7), 1401-34.



- Lee, S. & Dallman, M. (2008). Engaging in a reflective examination about diversity: Interviews with three pre-service teachers. *Multicultural Education*, 15(4), 36-44.
- Leung, C., & Franson, C. (2001). Mainstreaming: ESL as a diffused curriculum concern. In B. Mohan, C. Leung, & Davison, C. (Eds). (2001). *English as a second language in the mainstream: Teaching, learning and identity*. New York, NY: Longman-Pearson.
- Leung, C., & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: assessment as discourse and assessment of discourse. *Language Testing*, 21(3), 335-359.
- Lewis, A. (2004). Undoable goals. *The Education Digest*: 69(7), 68-9.
- Llosa, L. (2005). Assessing English learners' language proficiency: A qualitative investigation of teachers' interpretations of the California ELD standards. *The CATesol Journal*, 17(1), 7-18.
- Llosa, L. (2007). Validating a standards-based classroom assessment of English proficiency: A multitrait-multimethod approach. *Language Testing*, 24(4), 489-515.
- Llosa, L. (2008). Building and supporting a validity argument for a standards-based classroom assessment of English proficiency based on teacher judgments. *Educational Measurement: Issues and Practice*. 47(3), 32-42.
- Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, 78(2), 333-368.
- Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educational Assessment*, 14, 160-79.
- Marzano, R., Pickering, D., & McTighe, J. (1993). *Assessing student outcomes: performance assessment using the dimensions of learning model*. Association for supervision and curriculum development: Alexandria, VA.
- McMillan, J. (2003). Understanding and improving teachers' classroom assessment decision making implications for theory and practice. *Educational Measurement: Issues and Practice*. 22(4), 34-43.
- McMillan, J. (2004). *Classroom assessment: Principles and practice for effective instruction*. Boston, MA: Allyn and Bacon.



- McMillan, J. (2007). Formative classroom assessment: The key to improving student achievement. In J. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 1-7). New York, NY: Teachers College Press.
- McMillan, J., Myron, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research*, 95(4), 203-213.
- McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing* 18(4), 333-349.
- Meisels, S., Bickel, D., Nicholson, J., Xue, Y., & Atkins-Burnett, S. (2001). Trusting teachers' judgments: A validity study of a curriculum-embedded performance assessment in kindergarten to grade 3. *American Educational Research Journal*: 38(1), 73-95.
- Menken, K. (2006). Teaching to the test: How No Child Left Behind impacts language policy, curriculum, and instruction of English language learners. *Bilingual Research Journal*, 30(2), 521-46.
- Menken, K. (2008). *English language learners left behind*. New York, NY: Multilingual Matters.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18(2), 5-11.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing* 13(3), 241-56.
- Miller, K., & Satchwell, C. (2006). The effect of beliefs about literacy on teacher and student expectations: A further education perspective. *Journal of Vocational Education and Training*, 58(2), 135-150.
- Miller, E. R., Okum, I., Sinai, R., & Miller, K. S. (1999). *A study of the English language readiness of limited English proficient students to participate in New Jersey's statewide assessment system*. Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal.
- Mohan, B., Leung, C., & Davison, C. (Eds.). (2001). *English as a second language in the mainstream: Teaching, learning and identity*. New York, NY: Longman-Pearson.
- Mulvenon, S., Stegman, C. & Ritter, G. (2005). Test anxiety: A multifaceted study on the perceptions of teachers, principals, counselors, students, and parents. *International Journal of Testing*, 5(1), 37-61.
- National Clearinghouse for English Language Acquisition. (2006). *How many school aged Limited English Proficient (LEP) students are there in the US?: Frequently Asked*



- Questions*, Washington DC. Retrieved February 2, 2010, from <http://www.ncela.gwu.edu/faqs/>.
- National Clearinghouse for English Language Acquisition. (2009a). *The growing numbers of Limited English Proficient students, 1995/6-2005/6*. poster. Washington DC. Retrieved November 25, 2012, from http://www.ncela.gwu.edu/files/uploads/9/growingLEP_0809.pdf.
- National Center for Educational Statistics (2008). *Characteristics of the 100 largest public elementary and secondary school districts in the United States, statistical analysis report*. Washington DC. Retrieved February 2, 2010, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/3e/09/ac.pdf.
- National Center for Educational Statistics (2008a). *Identification of rural locales*. Washington DC. Retrieved April 30, 2010 from http://nces.ed.gov/ccd/rural_locales.asp.
- National Center for Educational Statistics (2010a). *Number of public school students and number and percentage of public school students who were English language learners (ELLs) by state: Selected school years 2000-1 through 2009-10*. Retrieved November 24, 2012 from <http://nces.ed.gov/programs/coe/tables/table-ell-1.asp>.
- National Center for Educational Statistics (2012). *State educational data profiles, 2010-11*. Washington DC. Retrieved November 24, 2012, from <http://nces.ed.gov/programs/stateprofiles/sresult.asp?mode=full&displaycat=1&s1=42>.
- New York State Education Department (2013). *OBE-FLS. Bilingual education and foreign language studies: Frequently Asked Questions, Assessment and testing of LEP/ELL Students*. Retrieved January 18, 2012, from <http://www.p12.nysed.gov/biling/bilinged/faq.html#testing>.
- Nieto, S. (2002). *Language, culture and teaching*. Mahwah, NJ: Erlbaum.
- Oakes, J. (2005). *Keeping track: How schools structure inequality*. London, UK: Yale University Press.
- Oh, S. (2001). Two types of input modification and EFL reading comprehension: Simplification vs. elaboration. *TESOL Quarterly*, 35(1), 69-96.
- Olson, K. (2007). Lost opportunities to learn: The effects of education policy on primary language instruction for English learners. *Linguistics and Education*, 18, 121-141.



- Osterlind, S., Miao, D., Sheng, Y., & Chia, R. (2004). Adapting item format for cultural effects in translated tests: Cultural effects on construct validity of the Chinese versions of the MBTI. *International Journal of Testing*, 4(1), 61-73.
- PARCC (2013). *Career and college readiness for the next generation*. Retrieved January 25, 2013 from [http://www.parcconline.org/sites/parcc/files/PARCC-Overview Oct2012.pdf](http://www.parcconline.org/sites/parcc/files/PARCC-Overview%20Oct2012.pdf).
- Payan, R. & Nettles, M. (2008). *Current state of English-language learners in the U.S. K-12 student population*. Princeton, NJ: Educational Testing Service. Retrieved March 31, 2010, from http://www.ets.org/Media/Conferences_and_Events/pdf/ELLSympisium/ELL_factsheet.pdf.
- Pedulla, J., Abrams, L., Maduas, G., Russell, M., Ramos, M., & Miao, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*. Chestnut Hill, MA. Boston College, national board on educational testing and public policy.
- Penfield, J. (1987). The regular classroom teacher's perspective. *TESOL Quarterly*, 21(1), 21-39.
- Pennsylvania Department of Education (2009a). *PSSA testing calendar best practice & recommendations*. Retrieved February 3, 2010, from <http://www.portal.state.us.pa>.
- Pennsylvania Department of Education (2009b). *PSSA and PSSA-M Accommodations guidelines for English language learners*. Retrieved February 3, 2010 from <http://www.portal.state.us.pa>.
- Pennsylvania Department of Education (2009c). *Educating students with limited English proficiency and English language learners*. Retrieved March 31, 2010 from http://www.portal.state.pa.us/portal/server.pt/community/pa_codes/7501/educating_students_with_limited_english_proficiency.
- Pennsylvania Department of Education (2012). *2012 Accommodations guidelines for English language learners*. Retrieved January 3, 2012 from http://www.education.state.pa.us/portal/server.pt/community/testing_accommodations__security/7448.
- Peregoy, S. & Boyle, O. (2005). *Reading, writing, and learning in ESL: A Resource book for teaching K-12 English learners* (4th ed). Boston, MA: Pearson Education.
- Pettit, S. K. (2011). Teachers' beliefs about English language learners in the mainstream classroom: A review of the literature. *International Multilingual Research Journal*, 5(2), 143-7.



- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18(4), 429-62.
- Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21(3), 249–258.
- Rea-Dickins, P., & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*, 17(2), 215-243.
- Reeves, J. (2004). “Like everybody else”: Equalizing educational opportunity for English language learners. *TESOL Quarterly*, 38(1), 43-66.
- Reiser, L. (2006). *English language learners in Pennsylvania schools*, Education Law Center. Retrieved February 2, 2010, from <http://www.elc-pa.org>.
- Rinaldi, C., & Samson, J. (2008). English language learners and response to intervention: Referral considerations. *Teaching Exceptional Children*, 40(5), 6-14.
- Rivera, C. & Collum, E. (2006). *State assessment policy and practice for English language learners*. Mahwah, NJ: Lawrence Erlbaum.
- Rivera, C. & Collum, E. (2006a). Including and accounting for English language learners in state assessments. In C. Rivera, & E. Collum (Eds.), *A National Review of State Assessment Policy and Practice for English Language Learners* (pp. xxxiii-xivii). Mahwah, NJ: Lawrence Erlbaum Associates.
- Rivera, C., Collum, E., Shafer, Willner, L., & Sia Jr., J. K. (2006). An analysis of state assessment policies addressing the accommodation of English language learners. In C. Rivera, & E. Collum (Eds.), *A national review of state assessment policy and practice for English language learners* (pp. 1-173). Mahwah, NJ: Lawrence Erlbaum Associates.
- Rivera, C., & Stansfield, C (2004). The effect of linguistic simplification of science test items on scores comparability. *Educational Assessment*, 9(3&4), 79-105.
- 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: George Washington University, Center for Equity and Excellence in Education.
- Rubin, H. & Rubin, I. (2005). *Qualitative interviewing: The art of hearing data*. 2nd edition, Thousand Oaks, CA: Sage Publications.
- Scriven, M. (1967). The methodology of evaluation. In R.W. Tyler, R.M. Gagne, and M. Scriven (Eds.), *Perspectives of curriculum evaluation*. Chicago, IL: Rand McNally.



- Seidman, I. (2006). *Interviewing as qualitative research*. New York, NY: Teachers College Press.
- Sharkey, J. & Layzer, C. (2000). Whose definition of success? Identifying factors that affect English language learners' access to academic success and resources. *TESOL Quarterly*, 34(1), 352-368.
- Shaver, A., Cuevas, P., Lee, O., Avalos, M. (2007). Teachers' perceptions of policy influences on science instruction with culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching*. 44(5), 725-46.
- Shohamy, E. (1997). Testing methods, testing consequences: Are they ethical? Are they fair? *Language Testing*, 14(3), 340-349.
- Sinaiko, H., & Brislin, R. (1973). Evaluating language translations: Experiments on three assessment methods. *Journal of Applied Psychology*, 57(3), 328-334.
- Sireci, S., Li, S., & Scarpati, S. (2003). The effect of test accommodation on test performance: A review of the literature. (Research Report 495). Amherst: University of Massachusetts School of Education, Center for Educational Assessment.
- Sireci, S., Scarpati, S., & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research*, 75(4), 457-490.
- Smarter Balanced Assessment Consortium (2013). *School years*. Retrieved January 25, 2013 from <http://www.smarterbalanced.org/timeline>.
- Stansfield, C. (2003). Test translation and adaptation in public education in the USA. *Language Testing*, 20(2), 189-207.
- Stansfield, C., & Bowles, M. (2006). Test translation and state assessment policies for English language learners. In C. Rivera, & E. Collum (Eds.), *State assessment policy and practice for English language learners* (pp. 175-314). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stansfield, C., & Rivera, C. (2001). Test accommodations for LEP students. *Eric Digest*, 1-6.
- Stiggins, R. J. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan*, 81(3), 191-201.
- Stiggins, R.J. (2007). Conquering the formative assessment frontier. In J. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp.8-28). New York, NY: Teachers College Press.



- Stiggins, R. J. (2007a). Assessment through the students' eyes. *Educational Leadership*, 64(8), 22-26.
- Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Albany, NY: SUNY Press.
- Tabors, P. (1997). *One child, two languages: A guide for preschool educators of children learning English as a second language*. Baltimore, MD: Brookes Publishing.
- Taylor, H., Anselmo, M., Foreman, A., Schatschneider, C., & Angelopoulos, J. (2000). Utility of Kindergarten teacher judgments identifying early learning problems. *Journal of Learning Disabilities*, 33(2), 200-210.
- Texas Department of Education (2008). *Texas student assessment program, 2008-9 accommodations manual*. Texas Education Agency, p. 18. Retrieved February 1, 2010, from http://ritter.tea.state.tx.us/student.assessment/admin/AccommManual_2008_09.pdf.
- Texas Department of Education (2009). *Grade 3-8 and 10, linguistically accommodated testing (LAT) test administrator manual for 2009*. Texas Education Agency. Retrieved February 2, 2010, from http://ritter.tea.state.tx.us/student.assessment/admin/rpte/LAT_Manual_2009.pdf.
- Thomas, W., & Collier, V. (1997). *School effectiveness for language minority students*. Alexandria, VA: National Clearinghouse for Bilingual Education.
- Timperley, H., & Phillips, G. (2003). Changing and sustaining teachers' expectations through professional development in literacy. *Teaching and Teacher Education*, 19(6), 627-641.
- Tindal, G., McDonald, M., Tadesco, M., & Glasgow, A. (2003). Alternate assessments in reading and math: Development and validation for students with significant disabilities. *Exceptional Children*. 69(4), 481-494.
- US Department of Education. (2004). *No Child Left Behind: Testing: Frequently Asked Questions*. Washington, DC. Retrieved January 25, 2010, from <http://www.ed.gov/nclb/accountability/ayp/testing-faq.html#4>.
- Valenzuela, A. (1999). *Subtractive schooling: U.S.-Mexican youth and the politics of caring*. Albany, NY: State University of New York Press.
- Van de Vijver, F., & Hambleton, R. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1(2), 89-99.



- Walker, A., Shafer, J., & Iiams, M. (2004). "Not in my classroom": Teacher attitudes towards English language learners in the mainstream classroom. *National Association for Bilingual Education Journal of Research and Practice*, 2, 130–160.
- Webb, J., & Aspey, S. (2003, June 10). *President Bush, Secretary Paige celebrate approval accountability plan under No Child Left Behind*. Retrieved January 6, 2004, from <http://www.ed.gov/news/pressreleases/2003/06/06102003.html>.
- WIDA (2011). *Access for ELLs*. Retrieved February 25, 2013 from <http://www.wida.us/assessment/ACCESS/>
- Willingham, W., Pollack, J., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal for Educational Measurement*, 39(1), 1-37.
- Willner, L.S., Rivera, C. & Acosta, B.D. (2008). *Descriptive study of state assessment policies for accommodating English language learners*. Arlington, VA: The George Washington University Center for Equity and Excellence in Education.
- Willner, L.S., Rivera, C., & Acosta, E.D. (2009). Ensuring accommodations used on content assessments are responsive to English language learners. *Reading Teacher*, 62(8), 696-8.
- Wilson, R., & Kittleson, J. (2012). The role of struggle in pre-service elementary teachers' experience as students and approaches to facilitating science learning. *Research in Science Education*, 42(4), 709-728.
- Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis and interpretation*. Thousand Oaks, CA: Sage Publications.
- Wolf, M. K., & Leon, S. (2009). An investigation of the language demands in content assessment for English language learners. *Educational Assessment*, 14(3), 139-59.
- Wyoming Department of Education ((2010). *Wyoming statewide assessment system, 2010 PAWS standard accommodations*. Retrieved February 3, 2010, from http://www.k12.wy.us/SA/Paws/Paws_Info/2009_10/PAWS_acomodations.PDF.
- Yilmaz-Tuzun, O. (2008). Pre-service elementary teachers' beliefs about science teaching. *Journal of Science Teacher Education*, 19, 183-204.
- Yoon, B. (2008). Uninvited guests: The influence of teachers' roles and pedagogies on the positioning of English. *American Educational Research Journal*, 45(2), 495-522.



- Young, J., Cho, Y., Ling, G., Cline, F., Steinberg, J., & Stone, E. (2008). Validity and fairness of state standards based assessments for English language learners. *Educational Assessment, 13*, 170-192.
- Youngs, C., & Youngs, G. (2001). Predictors of mainstream teachers' attitudes toward ESL students. *TESOL Quarterly, 35*(1), 97-120.
- Zimmerman, B., & Dibenedetto, M (2008). Mastery learning and assessment: Implications for students and teachers in an era of high-stakes testing. *Psychology in the Schools, 45*(3), 206-16.