**ASSESSING YOUNG LEARNERS: SELECTED REFERENCES**

**(Last updated 30 June 2020)**

Abedi, J. (2007). English language proficiency assessment and accountability. In Abedi, J. (Ed.), *English language proficiency assessment in the nation: Current status and future practice* (pp. 3-10). Berkley, CA: The Regents of the University of California.

Ai, P. T. N., Nhu, N. V. Q., & Thuy, N. H. H. (2019). Vietnamese EFL teachers’ classroom assessment practice at the implementation of the pilot primary curriculum. *International Journal of Language and Linguistics*, *7*(4), 172-177.

Arjoudis, S., & O’Loughlin, S. (2004). Tensions between validity and outcomes: Teacher assessment of written work of recently arrived immigrant ESL students. *Language Testing, 21*(3), 284-304.

Bailey, A., & Butler, F. (2004). Ethical considerations in the assessment of the language and content knowledge of U.S. school-age English language learners. *Language Assessment Quarterly, 1*(2&3), 177-193*.*

Bailey, A. L., & Heritage, M. (2008). *Formative assessment for literacy grades K-6: Building reading and academic language skills across the curriculum* (pp. 15-18). Thousand Oaks, CA: Corwin Press.

Baron, P. A., & Papageorgiou, S. (2016). *Setting language proficiency score requirements for English-as-a-Second-Language placement decisions in secondary education* (Research Report RR-16-17). Princeton, NJ: Educational Testing Service.

Beck, S. W.,Llosa, L., Black, K. & Trzeszkowski-Giese, A. (2015). Beyond the rubric: Think-alouds as a diagnostic assessment tool for high school writing teachers. *Journal of Adolescent and Adult Literacy* 58 (8): 668-679.

Biggar, H. (2005). NAEYC recommendations on screening and assessment of young English-language learners. *Young Children, 60*(66), 44-46.

Borsato, G. N. & Padilla, A. M. (2007). Educational assessment of English language learners. In L. A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment* (3rd ed., pp. 471-489). San Francisco, CA: Jossey Bass Publishers.

Brindley, G. (1998). Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing, 15*(1), 45–85.

Butler, Y. G. (2017). The role of affect in intra-individual variability in task performance for young learners. *TESOL Quarterly, 51*(3), 728-737.

Butler, F. A. & Stevens, R. (2001). Standardized assessment of the content knowledge of English language learners K-12: Current trends and old dilemmas. *Language Testing,* *18*(1), 409-427.

Butler, F. A., & Stevens, R. (1997). Oral language assessment in the classroom. *Theory into Practice,* *36*(4), 214-219.

Butler, Y. G., & Lee, J. (2006). On-task verus off-task self-assessments among Korean elementary school students studying English. *Modern Language Journal, 90*(4), 506-518.

Butler, Y.G., & Lee J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, *27*(1), 854-867. DOI: 10.1177/0265532209346370.

Canale, M., & McSwain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1-47.

Carless, D., & Lam, R. (2014). The examined life: Perspectives of lower primary school students in Hong Kong. *Education 3-13: International Journal of Primary, Elementary, and Early Years Education, 42*, 313-329.

Carpenter, K., Fuji, N., & Kataoka, H. (1995). An oral interview procedure for assessing second language abilities in children. *Language Testing, 12*(2), 157-81.

Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors’ classroom assessment practices: Purposes, methods, and procedures. *Language Testing, 21*(3), 360-389.

Cho, Y., Ginsburgh, M., Morgan, R., Moulder, B., Xi, X., & Hauck, M. C. (2016). *Designing the TOEFL® Primary™ tests* (ETS RM-16-02).Princeton, NJ: Educational Testing Service.

Cho, Y., & So, Y. (2014). *Construct-irrelevant factors influencing young English as a foreign language (EFL) learners' perceptions of test task difficulty* (Research Memorandum No. RM-14-04). Princeton, NJ: Educational Testing Service.

Chu, C. S. (2017). Translanguaging in reading comprehension assessment: Implications on assessing literal, inferential, and evaluative comprehension among ESL elementary students in Taiwan. *NYS TESOL Journal, 4*(2), 20-34.

Clark-Gareca, B. (2016). Elementary English language learners and classroom content tests. In M. A. Christison & N. Saville (Eds.), *Advancing the field of language assessment: Papers from TIRF Doctoral Dissertation Grantees. Studies in Language Testing 46* (pp. 131-147). Cambridge, UK: Cambridge University Press.

Coombe, C., & Davidson, P. (2012). Assessing young language learners: Issues, principles and practices. In H. Emery & F. Gardiner-Hyland (Eds.), *Contextualizing EFL for young learners: International perspectives on policy, practice and procedure* (pp. 283-296). Dubai, UAE: TESOL Arabia.

Dalton, S. (1979). Validation of the Language Assessment Scales. *Educational and Psychological Measurement, 39*, 1001-1003.

Davison, C. (2004). The contradictory culture of teacher-based assessment practices in Australian and Hong Kong secondary schools. *Language Testing, 21*(3), 305-334.

Douglas, D., (1998). Review of *Assessment for equity and inclusion: Embracing all our children* by A. L. Goodwin (Ed.). International Journal of Bilingual Education and Bilingualism, 1(2), 149-153.

Edelenbos, P. & Kubanek-German, A. (2004). Teacher assessment: The concept of ‘diagnostic competence’. *Language Testing, 21*(3), 259-283.

Elder, C., Iwashita, N. & McNamara, T. (2002). Estimating the difficulty of oral proficiency tasks: What does the test-taker have to offer? *Language Testing,* *19*(4), 347-368.

Espinosa, L. M. (2005). Curriculum and assessment considerations for young children from culturally, linguistically, and economically diverse backgrounds. *Psychology in the Schools, 42*(8*)*, 837-853.

Esquinca, A., Yaden, D., & Rueda, R. (2005). Current language proficiency tests and their implications for preschool English language learners. In J. Cohen, K. McAlister, K. Rolstad, & J. MacSwan (Eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism* (pp. 674-680). Somerville, MA: Cascadilla Press.

Evanini, K., Heilman, M., Wang, X., & Blanchard, D. (2015). *Automated scoring for the TOEFL Junior® Comprehensive Writing and Speaking Test* (Research Report RR-15-09). Princeton, NJ: Educational Testing Service.

Fulcher, G. (1996). Testing tasks: Issues in group design and the group oral. *Language Testing*, *13*(1), 23-51.

Fulcher, G. (1997). An English language placement test: Issues in reliability and validity. *Language Testing, 14*(2), 113-139.

Fulcher, G., & Reiter, R. M. (2003). Task difficulty in speaking tests. *Language Testing,* *20*(3), 321-344.

Gao, L., & Watkins, D.A. (2001). Identifying and assessing the conceptions of teaching of secondary school physics teachers in China. *The British Journal of Educational Psychology*, *71* (3), 443-469.

Geva, E. (2000). Issues in the assessment of reading disabilities in L2 children – Beliefs and research evidence. *Dyslexia, 6,* 13-28.

Gu, L., & So, Y. (2017). Strategies used by young English learners in an assessment context. In M. K. Wolf & Y. G. Butler (Eds.), *English language proficiency assessments for young learners* (pp. 118–135). New York, NY: Routledge.

Gunning, P., & Turner, C. E. (2018). Young language learners in classroom contexts: The development of strategy assessment methods and tools. In R. L. Oxford & C. M. Amerstorfer (Eds.), *Language learning strategies and individual learner characteristics: Situating strategy use in diverse contexts* (pp. 261-286). London, UK: Bloomsbury Academic.

Hasselgreen, A. (2005). Assessing the language of young learners. *Language Testing, 22*(3), 337-354.

Hasselgren, A. (2000). The assessment of English ability of young learners in Norwegian schools: An innovative approach. *Language Testing, 17*(2), 261-77.

Hsieh, C.-N. (2016). Examining content representativeness of a young learner language assessment: EFL teachers' perspectives. In M Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives* (pp. 93-107). London, UK: Springer.

Hsieh, C.-N., & Gu, L. (2020). Young language learners’ strategy use and perceptions of picture-based speaking tasks. In R. M. Damerow & K. M. Bailey (Eds.), *Chinese-speaking learners of English: Research, theory, and practice* (pp. 171-182). New York, NY: Routledge.

Hui, S. K. F., Brown, G. T., & Chan, S. W. M. (2017). Assessment for learning and for accountability in classrooms: The experience of four Hong Kong primary school curriculum leaders. *Asia Pacific Education Review*, *18*(1), 41-51.

Jeynes, W. H. (2006). Standardized tests and Froebel’s original kindergarten model. *Teacher’s College Record,* *108*(10), 1937-1959.

Johnstone, R. (2000). Context-sensitive assessment of modern languages in primary (elementary) and secondary education: Scotland and the European experience. *Language Testing, 17*(2), 123-143.

Kondo-Brown, K. (2004). Investigating interviewer-candidate interactions during oral interviews for child L2 learners. *Foreign Language Annals, 37*(4), 602-615.

Lalani, S. S., & Rodrigues, S. (2012). A teacher's perception and practice of assessing the reading skills of young learners: A study from Pakistan. *Journal on English Language Teaching*, *2*(4), 23-33.

Lee, S., & Winke, P. (2018). Young learners’ response processes when taking computerized tasks for speaking assessment. *Language Testing, 35*(2), 239-269.

Leung, C. (2005). Classroom teacher assessment of second language development: Construct as practice. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 25-43). New York, NY: Routledge.

Lester, J. N., & Anders, A. D. (2014). Complicating translation: Children with refugee status and special education testing. *NYS TESOL Journal, 1*(2), 25-38.

Limbos, M. M. & Geva, G. (2001). Accuracy of teacher assessments of second-language students at risk for reading disability. *Journal of Learning Disabilities, 34*, 136-151.

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*(1), 489-515.

Llosa, L. (2008). Building and supporting a validity argument for standards-based classroom assessment of English proficiency based on teacher judgments. *Educational Measurement: Issues and Practice, 27(3), 32-42.*

Llosa, L. & Slayton, J. (2009). Using program evaluation to inform and improve the education of young English learners in U.S. schools. *Language Teaching Research, 13* (1), 35-54.

Maerten-Rivera, J., Huggins-Manley, A.C., Adamson, K., Lee, O., &Llosa, L. (2015). Development and validation of a measure of elementary teachers’ science content knowledge in two multi-year teacher professional development intervention projects. *Journal of Research in Science Teaching* 52 (3): 371-396.

Maxwell, L. A. (2013). Common core ratchets up language demands for language-learners. *Education Week, 33*(10), 14-16.

McKay, P. (2005). Research into the assessment of school-age language learners. *Annual Review of Applied Linguistics*, *25*, 243-263.

McKay, P. (2006). *Assessing young language learners*. New York, NY: Cambridge University Press.

Mercado, L. A. (2015). Standards-based assessment for young learners. In C. N. Giannikas, L. McLoughlin, & G. Fanning (Eds.), *Children learning English: From research to practice* (pp. 167-185). Reading, UK: Garnet Education.

New York State Education Department. (2007). *New York State Testing Program NYSESAT Sampler Grades K-1*. Retrieved From: http://www.p12.nysed.gov/assessment/nyseslat/samplers/2007/

New York State Education Department. (2012). *2012 New York State English as a Second Language Achievement Test (NYSESLAT) school administrator's manual*. Retrieved from http://www.p12.nysed.gov/assessment/nyseslat/archive.html

New York State Education Department. (2013). *New York State Testing Program NYSESLAT: Guide to the 2013 NYSESLAT.* Retrieved from: http://www.p12.nysed.gov/assessment/nyseslat/

Newton, X., & Llosa, L.(2010). Towards a more nuanced approach to program effectiveness assessment: Hierarchical linear models (HLM) in K-12 program evaluation. *American Journal of Evaluation*, *31*(2), 162-179.

Nikolov, M. (2016). Trends, issues, and challenges in assessing young language learners. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives* (pp. 1-17). London, UK: Springer.

O’Sullivan, B. (2000). Exploring gender and oral proficiency interview performance. *System, 28,* 373-386.

Pellegrini, A. D. (1998). Play and the assessment of young children. In O. N. Saracho & B. Spodek (Eds.), *Multiple perspectives on play in early childhood education* (pp. 220-239). Albany, NY: State University of New York Press.

Pham, L. A. (2013). A case study into English classroom assessment practices in three primary schools in Hanoi: Implications for developing a contextualized formative assessment practice framework. *VNU Journal of Foreign Studies*, *29*(1), 1-16.

Phillips, J., & Draper, J. (1994). National standards and assessments: What does it mean for the study of second languages in the schools? In G.K. Crouse (Ed.), *Meeting new challenges in the foreign language classroom* (pp. 1-8). Lincolnwood, IL: National Textbook.

Porter, S. G., & Vega, J. (2007). Overview of existing English language proficiency tests. In Abedi, J. (Ed.). *English language proficiency assessment in the nation: Current status and future practice* (pp. 93-102) Berkley, CA: The Regents of the University of California.

Rea-Dickins, P. (2000). Assessment in early years language learning contexts. *Language Testing, 17*(2)115-122.

Rea-Dickins, P. (2000). Current research and professional practice: Reports of work in progress in the assessment of young language learners. *Language Testing, 17*(2), 245-249.

Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing, 18*(4), 429-462.

Rea-Dickins, P., & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing, 17*(2), 215-243.

Roach, A. T., McGrath, D., Wixson, C., & Talapatra, D. (2010). Aligning an early childhood assessment to state kindergarten content standards: Application of a nationally recognized alignment framework. *Educational Measurement: Issues and Practice, 29*(1), 25-37.

Russell, R. L., & Grizzle, K. L. (2008). Assessing child and adolescent pragmatic competencies: Towards evidence-based assessment. *Clinical Child and Family Psychology Review*, *11*, 59-73.

Sanchez, L. (2006). Bilingualism/second-language research and the assessment of oral proficiency in minority bilingual children. *Language Assessment Quarterly, 3*(2), 117-149.

Schappe, J. F. (2005). Early childhood assessment: A correlational study of the relationship among student performance, student feelings, and teacher perceptions. *Early Childhood Education Journal, 33*(3), 187-193.

Tan, K. H. (2013). Variation in teachers’ conceptions of alternative assessment in Singapore primary schools. *Educational Research for Policy and Practice*, *12*(1), 21-41.

Tan, L. L., Wigglesworth, G., & Storch, N. (2011). Naplan language assessments for Indigenous children in remote communities, pair interactions and mode of communication. *Austra-lian Review of Applied Linguistics, 33*(3), 1–24.

Upshur, J. (1967). Testing foreign-language function in children. *TESOL Quarterly, 4*, 31-34.

Weir, C. (2005). *Language testing and validation: An evidence-based approach*. Basingstoke, UK: Palgrave McMillan.

WIDA Consortium. (2008). *ACCESS for ELLs listening, speaking, writing, and reading sample items 2008: Grade Kindergarten*. Wisconsin: Board of Regents of the University of Wisconsin System.

WIDA Consortium. (2008). *WIDA MODEL Measure of Developing English Language student response booklet: Grade K* (version 1.0). Wisconsin: Board of Regents of the University of Wisconsin System.

WIDA Consortium. (2008). *WIDA MODEL Measure of Developing English Language test administration manual: Grade K* (version 1.0). Madison, WI: Board of Regents of the University of Wisconsin System.

WIDA Consortium. (2008). *WIDA MODEL Measure of Developing English Language test administrator script: Grade K* (version 1.0). Madison, WI: Board of Regents of the University of Wisconsin System.

Wigglesworth, G., Simpson, J., & Loakes, D. (2011). Naplan language assessments for Indigenous children in remote communities. *Australian Review of Applied Linguistics*, *34*(3), 320-343.

Wiliam, D., & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal, 22*(5), 537-548.

Winke, P., Lee, S., Ahn, J. I., Choi, I., & Yoon, H. J. (2018). The cognitive validity of child English language tests: What young language learners and their native-speaking peers can reveal. *TESOL Quarterly, 52*(2), 274-303.

Wolf, M. K., & Butler, Y. G. (Eds.), (2017). *English language proficiency assessments for young learners*. New York, NY: Routledge.

Wu, W. M., & Stansfield, C. W. (2001). Towards authenticity of task in test development. *Language Testing, 18*(2), 187-206.

Yang, T. L. (2008). Factors affecting EFL teachers' use of multiple classroom assessment practices with young language learners. *English Teaching & Learning*, *32*(4), 85-123.

Yildirim, R., & Orsdemir, E. (2013). Performance tasks as alternative assessment for young EFL learners: Does practice match the curriculum proposal?. *International Online Journal of Educational Sciences*, *5*(3), 562-574.

Yin, M. (2010). Understanding classroom language assessment through teacher thinking research. *Language Assessment Quarterly, 7*(2), 175-194.

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.

Zehr, M. (2004). Tests of youngest English-learners spark controversy. *Education Week, 24*(12), 1-16.