

Evidence of the Validity of Teaching Strategies GOLD® Assessment Tool for English Language Learners and Children with Disabilities

Do-Hong Kim, Richard G. Lambert and Diane C. Burts

Research Findings: This study examined the measurement equivalence of the Teaching Strategies GOLD® assessment system across subgroups of children based on their primary language and disability status. This study is based on teacher-collected assessment data for 3-, 4-, and 5-year-old children for the fall of 2010, winter of 2010, and spring of 2011 checkpoint periods. Differential item functioning (DIF) analysis was conducted to determine whether items functioned differentially (a) between children with and without disabilities, (b) between English language learners (ELLs) and non-ELLs, and (c) between Spanish-speaking ELLs and non-ELLs. Results showed that the majority of items displayed little or no DIF. One item pertaining to the use of conventional grammar was consistently identified as having DIF and requires further review. **Practice or Policy:** Overall the findings suggest that the Teaching Strategies GOLD® is a viable observation-based, authentic assessment measure and is equally valid and reliable for children with disabilities and for those whose home language is not English. Implications for teacher education and training are discussed.

Citation:

Lambert, R. G., Kim, D. H., & Burts, D. C. (2013). Using teacher rating to track the growth and development of young children using the Teaching Strategies GOLD® assessment system. *Journal of Psychoeducational Assessment*, XX(X), 1-13. doi: 10.1177/0734282913485214

REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Atkins-Burnett, S. (2007). *Measuring children's progress from preschool through third grade*. Retrieved from <http://www.mathematica-mpr.com/publications/PDFs/measchildprogress.pdf>
- August, O., Carlo, M., Dressier, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research and Practice, 20*(1), 50—57.
- Ayoub, C. C., & Fischer, K. W. (2006). Developmental pathways and intersections among domains of development. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 62—81). Oxford, UK: Blackwell.
- Bagnato, S. J. (2005). The authentic alternative for assessment in early intervention: An emerging evidence-based practice. *Journal of Early Intervention, 28*(1), 17—22.
- Bagnato, S. J., McLean, M., Macy, M., & Neisworth, J. T. (2011). Identifying instructional targets for early childhood via authentic assessment: Alignment of professional standards and practice-based evidence. *Journal of Early Intervention, 33*, 243—253.
- Barnett, W. S., Epstein, D. J., Carolan, M. E., Fitzgerald, J., Ackerman, D. J., & Friedman, A. H. (2010). *The state of preschool 2010*. Retrieved from <http://nieer.org/yearbook/pdf/yearbook>
- Barton, L. R., Spiker, D., & Williamson, C. (2012). Characterizing disability in Head Start programs: Not so clearcut. *Early Childhood Research Quarterly, 27*, 596—612.
- Bordignon, C. M., & Lam, T. C. M. (2004). The early assessment conundrum: Lessons from the past, implications for the future. *Psychology in the Schools, 41*, 737—749.
- Brenneman, K. (2011). Assessment for preschool science learning and learning environments. *Early Childhood Research and Practice, 13*(1). Retrieved from <http://ecrp.uiuc.edu/v13n1/brenneman.html>
- Bridges, M., Cohen, S. R., McGuire, L. W., Yamada, H., Fuller, B., Mireles, L., & Scott, L. (2012). Bien educado: Measuring the social behaviors of Mexican American children. *Early Childhood Research Quarterly, 27*, 555—567.
- Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* 6th ed., (pp. 793—828). New York, NY: Wiley.
- Cabell, S. Q., Justice, L. M., Zucker, T. A., & Kilday, C. R. (2009). Validity of teacher report for assessing the emergent literacy skills of at-risk preschoolers. *Language and Speech & Hearing Services in Schools, 40*, 161—173.
- Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice, 17*, 31—44.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Erlbaum.

- Collins, M. F. (2010). ELL preschoolers' English vocabulary acquisition from storybook reading. *Early Childhood Research Quarterly*, 25, 84—97.
- Conboy, B. T., & Thal, D. J. (2006). Ties between the lexicon and grammar: Cross-sectional and longitudinal studies of bilingual toddlers. *Child Development*, 77, 712—735.
- Copple, C., & Bredekamp, S. (Eds.). (2009). *Developmentally appropriate practice in early Childhood programs serving children from birth to age 8* (3rd ed). Washington, DC: National Association for the Education of Young Children.
- Dhuey, E., & Lipscomb, S. (2010). Disabled or young? Relative age and special education diagnoses in schools. *Economics of Education Review*, 29, 857—872.
- Dichtelmiller, M. L., Jablon, I. R., Meisels, S., J., & Marsden, D. B. (2001). *The Work Sampling System for Head Start*. New York, NY: Pearson Early Learning.
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language. Young children at home and school*. Baltimore, MD: Brookes.
- Dorans, N. I., Schmitt, A. P., & Bleistein, C. A. (1992). The standardization approach to assessing comprehensive differential item functioning. *Journal of Educational Measurement*, 29, 309—319.
- Douglas, I. A., Roussos, L. A., & Stout, W. (1996). Item-bundle DIP hypothesis testing: Identifying suspect bundles and assessing their differential functioning. *Journal of Educational Measurement*, 33, 465—484.
- Downer, J. T., Lopez, M. L., Grimm, K. J., Hamagami, A., Pianta, R. C., & Howes, C. (2012). Observations of teacher-child interactions in classrooms serving Latinos and dual language learners: Applicability of the Classroom Assessment Scoring System in diverse settings. *Early Childhood Research Quarterly*, 27, 21—32.
- Early, D. M., Iruka, I. U., Ritchie, S., Barbarin, O. A., Winn, D-M. C., Crawford, G. M Pianta, R. C. (2010). How do pre-kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, 25, 177—193.
- Espinosa, L. (2010). *Getting it right for children from diverse backgrounds: Applying research to improve practice*. Upper Saddle River, NJ: Pearson Education.
- Fantsizzo, J., Hightower, D., Grim, S., & Montes, G. (2002). Generalization of the Child Observation Record: A validity study of diverse samples of urban, low-income preschool children. *Early Childhood Research Quarterly*, 17, 106—125.
- Feldman, E. N. (2010). Benchmarks curricular planning and assessment framework: Using standards without introducing standardization. *Early Childhood Education journal*, 38, 233—242.
- Finch, I. (2011). The use of multiple imputation for missing data in uniform DIF analysis: Power and Type I error rates. *Applied Measurement in Education*, 24, 281—301.
- Gallagher, P. A., & Lambert, R. G. (2006). Classroom quality, concentration of children with special needs, and child outcomes in Head Start. *Exceptional Children*, 73(1), 31—52.
- Garcia, E., & Jensen, B. (2009). Early educational opportunities for children of Hispanic origins. *SRCD Social Policy Report*, 23(2), 3—19.
- Goldstein, P. (2004). Helping young children with special needs develop vocabulary. *Early Education Journal*, 32, 1-43.
- Grishans-Brown, J., Hallam, R., & Brookshire, R. (2006). Using authentic assessment to

- evidence children's progress toward early learning standards. *Early Childhood Education Journal*, 34(1), 45—51.
- Grisham-Brown, J., Hallam, R. A., & Pretti-Frontczak, K. (2008). Preparing Head Start personnel to use a curriculum based assessment: An innovative practice in the 'age of accountability.' *Journal of Early Intervention*, 30(4), 271—281.
- Gullo, D. F. (2006). Assessment in kindergarten. In D. F. Gullo (Ed.), *K today: Teaching and learning in the kindergarten year* (pp. 138—147). Washington, DC: National Association for the Education of Young Children.
- Hallam, R., Grisham-Brown, J., Gao, X., & Brookshire, R. (2007). The effects of outcomes-driven authentic assessment on classroom quality. *Early Childhood Research and Practice*, 9(2). Retrieved from <http://ecrp.luiuc.edu/v9n2/hallam.html>
- Halle, T., Hair, E., Wandner, L., McNamara, M., & Chien, N. (2012). Predictors and outcomes of early versus later English language proficiency among English language learners. *Early Childhood Research Quarterly*, 27, 1—20.
- Harper, L. V., & McCluskey, K. S. (2002). Caregiver and peer responses to children with language and motor disabilities in inclusive preschool programs. *Early Childhood Research Quarterly*, 17, 148—166.
- Hernandez, D. J., Takanishi, R., & Marotz, K. G. (2009). Life circumstances and public policies for young children in immigrant families. *Early Childhood Research Quarterly*, 24, 487—501.
- Neroman, C., Burts, D., Berke, K., & Bickart, T. (2010). *Teaching Strategies GOLD™ objectives for development and learning*. Washington, DC: Teaching Strategies.
- HighScope Educational Research Foundation. (1992). *High/Scope Child Observation Record (COR) for ages 2 1/2-6*. Ypsilanti, MI: High/Scope Press.
- HighScope Educational Research Foundation. (2002). *Infant-Toddler COR: Appendix B: Development and validation*. Retrieved from <http://www.highscope.org/Content.asp?ContentId=85>
- Hindman, A. H., Skibbe, L. E., Miller, A., & Zimmerman, M. (2010). Ecological contexts and early learning: Contributions of child, family, and classroom factors during Head Start, to literacy and mathematics growth through first grade. *Early Childhood Research Quarterly*, 25, 235—250.
- Hirsh-Pasek, K., Kochanoff, A., Newcombe, N. S., & de Villiers, J. (2005). Using scientific knowledge to inform preschool assessment: Making a case for "empirical validity." *SRCD Social Policy Report*, 14(1), 3—19.
- Hom, E., & Kang, J. (2012). Supporting young children with multiple disabilities: What do we know and what do we still need to learn? *Topics in Early Childhood Special Education*, 31(4), 241—248.
- Hyson, M. (2008). *Enthusiastic and engaged learners: Approaches to learning in the early childhood classroom*. New York, NY: Teachers College Press.
- Justice, L. M., Cottone, E. A., Mashburn, A., & Rimm-Kaufman, S. E. (2008). Relationships between teachers and preschoolers who are at risk: Contribution of children's language skills, temperamentally based attributes, and gender. *Early Education & Development*, 19, 600—621.
- Keilty, B., LaRocco, D. J., & Casell, F. B. (2009). Early interventionists' reports of authentic

- assessment methods— through focus group research. *Topics in Early Childhood Special Education*, 28, 244—256.
- Kim, D.-H., Lambert, R. G., & Burts, D. C. (2012). Validating a developmental scale for young children using the Rasch model: Applicability of the Teaching Strategies GOLD® assessment system. Manuscript submitted for publication.
- Kim, D.-H., & Smith, J. D. (2010). Evaluation of two observational assessment systems for children's development and learning. *NHSA Dialog*, 13, 253—267.
- Klintwall, L., HoIm, A., Eriksson, M., Carisson, L. H., Olsson, M. B., Hedvall, A Femell, E. (2011). Sensory abnormalities in autism: A brief report. *Research in Developmental Disabilities*, 32(2), 795—800.
- Kong, N. Y., & Carta, J. J. (2011). Responsive interaction interventions for children with or at risk for developmental delays: A research synthesis. *Topics in Early Childhood Special Education*. Advance online publication. doi:10.1177/0271121411426486
- Laing, S. P., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech, & Hearing Services in Schools*, 34(1), 4-4—55.
- Lambert, R. G., Kim, D.-H., & Burts, D. C. (2012). The measurement properties of the Teaching Strategies GOLD® assessment system. Manuscript submitted for publication.
- Lambert, R. G., Kim, D.-H., & Burts, D. C. (in press). Using teacher ratings to track the growth and development of young children using the Teaching Strategies GOLD® assessment system. *Journal of Psychoeducational Assessment*.
- Lambert, R. G., Kim, D.-H., Taylor, H., & McGee, J. R. (2010). Technical manual for the Teaching Strategies GOLD® assessment system. University of North Carolina at Charlotte: Center for Educational Measurement and Evaluation.
- Lonigan, C. J., & Shanahan, T. (2008). Executive summary of the report of the National Early Literacy Panel. Retrieved from <http://www.nifl.gov/earlychildhood/NELP/NELPreport.html>
- Lopez, E. J., Salas, L., & Flores, J. P. (2005). Hispanic preschool children: What about assessment and intervention? *Young Children*, 60(6), 48—54.
- Lutz, M. N., Fantuzzo, J., & McDermott, P. (2002). Multidimensional assessment of emotional and behavioral adjustment problems of low-income preschool children: Development and initial validation. *Early Childhood Research Quarterly*, 17, 338—355.
- Macy, M., & Bagnato, S. J. (2010). Keeping it “R-E-A-L” with authentic assessment. *NI-ISA Dialog*, 13, 1—20.
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of National Cancer Institute*, 22, 719—748.
- McClelland, M. M., & Cameron, C. E. (2011). Self-regulation in early childhood: Improving conceptual clarity and developing ecologically valid measures. *Child Development Perspectives*, 6, 136—142. doi:10.1111/j.1750-8606.2011.00191.x
- Meisels, S. J., & Atkins-Burnett, S. (2005). *Developmental screening in early childhood: A guide* (5th ed). Washington, DC: National Association for the Education of Young Children.
- Meisels, S. J., Bickel, D. 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.
- Meisels, S. J., Liaw, F., Dorfman, A., & Nelson, R. F. (1995). *The Work Sampling System*:

- Reliability and validity of a performance assessment for young . Children. *Early Childhood Research Quarterly*, 10, 277—296.
- Meisels, S. J., Wen, X., & Beachy-Quick, K. (2010). Authentic assessment for infants and toddlers: Exploring the reliability and validity of the Ounce Scale. *Applied Developmental Science*, 14, 55—71.
- Meisels, S. J., Xue, Y., & Shambloott, M. (2008). Assessing language, literacy, and mathematics skills with Work Sampling for Head Start. *Early Education & Development*, 19, 963—981.
- Michaelides, M. P. (2008). An illustration of a Mantel-Haenszel procedure to flag misbehaving common items in test equating. *Practical Assessment Research & Evaluation*, 13(7). Retrieved from <http://pareonline.net/pdf/v13n7.pdf>
- Moreno, A. J., & Kiute, M. M. (2011). Infant-toddler teachers can successfully employ authentic assessment: The Learning Through Relating system. *Early Childhood Research Quarterly*, 26, 484—496. National Association for the Education of Young Children, & National Association of Early Childhood Specialists in
- State Departments of Education. (2003). Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8. Joint position statement. Retrieved from www.naeyc.org/positionstatements/capel National Association for the Education of Young Children, & National Association of Early Childhood Specialists in
- State Departments of Education. (2005). Screening and assessment of young English-language learners: Supplement to the NAEYC and NAECS/SDE joint position Statement on early childhood curriculum, assessment, and program evaluation. Retrieved from www.naeyc.org/positionstatements/cspe/
- Paez, M., & Rinaldi, C. (2006). Predicting English word reading skills for Spanish-speaking students in first grade. *Topics in Language Disorders*, 26, 338—350.
- Paez, M. M., Tabors, P. O., & Lopez, L. M. (2007). Dual language and literacy development of Spanish-speaking preschool children. *Journal of Applied Developmental Psychology*, 28(2), 85—102.
- Parish-Morris, J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M., & Tager-Flusberg, H. (2007). Children with autism illuminate the role of social intention in word learning. *Child Development*, 78, 1265-1287.
- Pena, E. D., & Halle, T. G. (2011). Assessing preschool dual language learners: Traveling a multiforked road. *Child Development Perspectives*, 5(1), 28—32.
- Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research*, 37, 239—253.
- Qi, C. H., Kaiser, A. P., & Milan, S. (2006). Children's behavior during teacher-directed and child-directed activities in Head Start. *Journal of Early Intervention*, 28(2), 97—110.
- Qi, C. H., & Marley, S. C. (2009). Differential item functioning analysis of the Preschool Language Scale-4 between English-speaking Hispanic and European American children from low-income families. *Topics in Early Childhood Special Education*, 29(3), 171—180.
- Ray, A., Bowman, B., & Brownell, J. O. (2006). Teacher-child relationships, socio-emotional development, and school achievement. In B. Bowman, & E. K. Moore (Edt.), *School*

- readiness and social-emotional development: Perspectives on cultural diversity, (pp. 7—22). Washington, DC: National Black Child Development Institute.
- Reardon, S. F., & Galindo, C. (2007). Patterns of Hispanic students' math skill proficiency in the early elementary grades. *Journal of Latinos & Education*, 6(3), 229—251.
- Reardon, S. F., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal*, 46, 853—891.
- Rodriguez, B. L., & Guiberson, M. (2011). Using a teacher rating scale of language and literacy skills with preschool children of English-speaking, Spanish-speaking, and bilingual backgrounds. *Early Childhood Education Journal*, 39, 303—311.
- Rothstein-Fisch, C., Trumbull, E., & Garcia, S. G. (2009). Making the implicit explicit: Supporting teachers to bridge cultures. *Early Childhood Research Quarterly*, 24, 474—486.
- Rudasill, K. M., Rimm-Kaufman, S. E., Justice, L. M., & Pence, K. (2006). Temperament and language skills as predictors of teacher-child relationship quality in preschool. *Early Education & Development*, 17, 271—291.
- Schweinhart, L. J., McNair, S., Barnes, H., & Lamer, M. (1993). Observing young children in action to assess their development: The High/Scope Child Observation Record Study. *Educational and Psychological Measurement*, 53, 445—455.
- Sekino, Y., & Fantuzzo, J. (2005). Validity of the Child Observation Record: An investigation of the relationship between COR dimensions and social-emotional and cognitive outcomes for Head Start children. *Journal of Psychoeducational Assessment*, 23, 242—261.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4—14.
- Snow, C. E., & Van Hemel, S. B. (Eds.). (2008). *Early childhood assessment: Why, what, and how?* Report of the National Research Council of the National Academies. Retrieved from <http://www.nap.edu/catalog/12446.html>
- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, 38, 934—947.
- Teaching Strategies, LLC. (2001). *Creative Curriculum Developmental Continuum for Ages 3—5*. Washington, DC: Author.
- Teaching Strategies, LLC. (2005). *Expanded forerunners of the Creative Curriculum Developmental Continuum for Ages 3—5*. Washington, DC: Author.
- Teaching Strategies, LLC. (2006). *Creative Curriculum Developmental Continuum for Infants, Toddlers, and Twos*. Washington, DC: Author.
- Uccelli, P., & Paez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools*, 38, 225—236. U.S. Department of Health and Human Services, Administration for Children and Families.
- Office of Head Start. (2010). *The Head Start child development and early learning framework: Promoting positive outcomes in early childhood programs serving children 3—5 years old*. Washington, DC: Author.
- Vitiello, V. E., Booren, L. M., Downer, I. T., & Williford, A. P. (2012). Variation in children's classroom engagement throughout a day in preschool: Relations to classroom and child factors. *Early Childhood Research Quarterly*, 27, 210—220.

- Waterman, C., McDermott, P. A., Fantuzzo, J. W., & Gadsden, V. L. (2012). The matter of assessor variance in early childhood education—Or whose score is it anyway? *Early Childhood Research Quarterly*, 27(1), 46—54. doi:10.1016/j.ecresq.2011.06.003
- Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research and Evaluation*, 12(3), 1—26.
- Yesil-Dagli, U. (2011). Predicting ELL students' beginning first grade English oral reading fluency from initial kindergarten vocabulary, letter naming, and phonological awareness skills. *Early Childhood Research Quarterly*, 26, 15—29.
- Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, 30, 233—251.
- Zwick, R., & Thayer, D. T. (1996). Evaluating the magnitude of differential item functioning in polytomous items. *Journal of Educational and Behavioral Statistics*, 21, 187—201.