

Current R-PAS® Transitional Child and Adolescent Norms

January 29, 2014; updated February 28, 2016

Summary authored by Gregory J. Meyer, Donald J. Viglione, and Luciano Giromini

With Data and Research Contributions from:

Janell Crow, Carla Hisatugo, Ana Cristina Resende, Daria Russo, and Jessica Swanson

And Scoring Help from:

Heidi Miller, Vanessa Laughter, and Andrew Williams

As documented in our Problems with Child and Adolescent Normative Data Reference Data document, the existing Comprehensive System (CS) norms for children and adolescents are seriously in error on a number of variables and this ends up making normal children look pathological. Our initial solution to this problem, as described in the Statement, was to use CS normative data that had been published in the 2007 JPA Supplement on CS International Reference Samples (Meyer, Erdberg, & Shaffer, 2007), where normative data was organized into three age ranges: 5 to 8, 9 to 12, and 13 to 18. Such large age bins are not optimal, however, and so we have been gathering additional child and adolescent data so more fine-grained age-based norms could be implemented. Although still "transitional" rather than final R-PAS norms, we now have enough data to replace the former CS-based norms with data for R-PAS. This technical page describes the samples and procedures we used to form the new transitional norms.

We started with new R-PAS administered protocols from Brazil ($n = 197$) and the US ($n = 113$) spanning the age range from 6 to 17. To this we added a small number of R-Optimized modeled records that had originally been collected using CS administration guidelines. These protocols encompassed ages 7 to 12 and were collected in the US ($n = 24$) or Italy ($n = 11$). All protocols were coded using R-PAS guidelines. In the combined sample each gender was approximately evenly represented (51.7% female, 45.1% male, 3.2% missing information).

There was variability in the number of cases available at each age, with ns by age as follows: 6 = 3, 7 = 56, 8 = 74, 9 = 60, 10 = 77, 11 = 43, 12 = 15, 13 = 11, 14 = 2, 15 = 1, 16 = 2, and 17 = 1. Because we want the data to generalize internationally, we tried to give each country relatively equal weight in the final analyses for each age level. To accomplish this, individual cases from a country at a particular age were assigned weights that ranged from a low of .64 to a high of 3.0. Weights were assigned in such a manner that the overall sample size at a given age would remain fixed. For instance, among the 74 8-year-old children, 22 were from the US, 50 were from Brazil, and 2 were from Italy. In order to more optimally equalize these samples, the US children were given a weight of 1.25, the Brazilian children were given a weight of .80, and the Italian children were given a weight of 3. Thus, in the weighted analyses, the effective sample sizes were: US = 28 ($22 * 1.25$ rounded), Brazil = 40 ($50 * .80$), and Italy = 6 ($2 * 3.0$), such that the overall sample size for 8-year-olds remained at 74.

We concluded that the individual ages from 7 to 12 had enough cases to analyze separately. However, we combined the cases from ages 13 to 17. We also retained the very small sample of 6-year-olds ($n = 3$), though given its size it made little contribution to the analyses discussed next. To anchor the adult end of the developmental continuum, the R-PAS normative sample of full-text records ($n = 118$) was subdivided into three age groups: young adults age 18 to 26 ($n = 21$, $M = 22.9$), middle adults age 27 to 54 ($n = 75$, $M = 38.6$), and older adults age 55 to 86 ($n = 21$, $M = 64.8$). The one 17-year-old in the R-PAS norms was placed in with the group of cases from 13 to 17, resulting in a sample of $n = 18$ with an average age of 13.9.

To generate normative expectations we used two statistical procedures to maximize our ability to detect genuine developmental changes from these small and imperfect samples. First, we relied on a procedure called "continuous norming," which uses curve fitting polynomial regression to estimate developmental trends and allows all ages to contribute to the normative expectations (Zachary & Gorsuch, 1985). As such, the number of actual protocols contributing to the regression estimates was 463. Continuous

norming has been embraced by test publishers as a means of enhancing accuracy in generating normative estimates, as it can be used to predict means and standard deviations, as well as skew and kurtosis. As pointed out by Zhu and Chen (2011), continuous norming has been used to develop norms for the Stanford-Binet Intelligence Scales – Fifth Edition (Roid, 2003), Wechsler Adult Intelligence Scale – Fourth Edition (Wechsler, 2008), Wechsler Individual Achievement Test – Third Edition (Wechsler, 2009), and Woodcock-Johnson III (Woodcock, McGrew, & Mather, 2001), among others. More recently, the continuous norming approach has been expanded to encompass “inferential norming,” which uses expert judgments about logical developmental trajectories to potentially adjust the polynomial curves and make a final determination about the most reasonable developmental trend, as well as to model complete score distributions (i.e., M, SD, skew, and ultimately percentiles) at all age levels (Zhu & Chen, 2011). In both a simulation study and using actual normative data from the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV; Wechsler, 2003) inferential norming with age-based samples of $n \geq 50$ produced norms that were as accurate as norms produced the traditional way with age-based samples of $n = 200$ (Zhu & Chen, 2011).

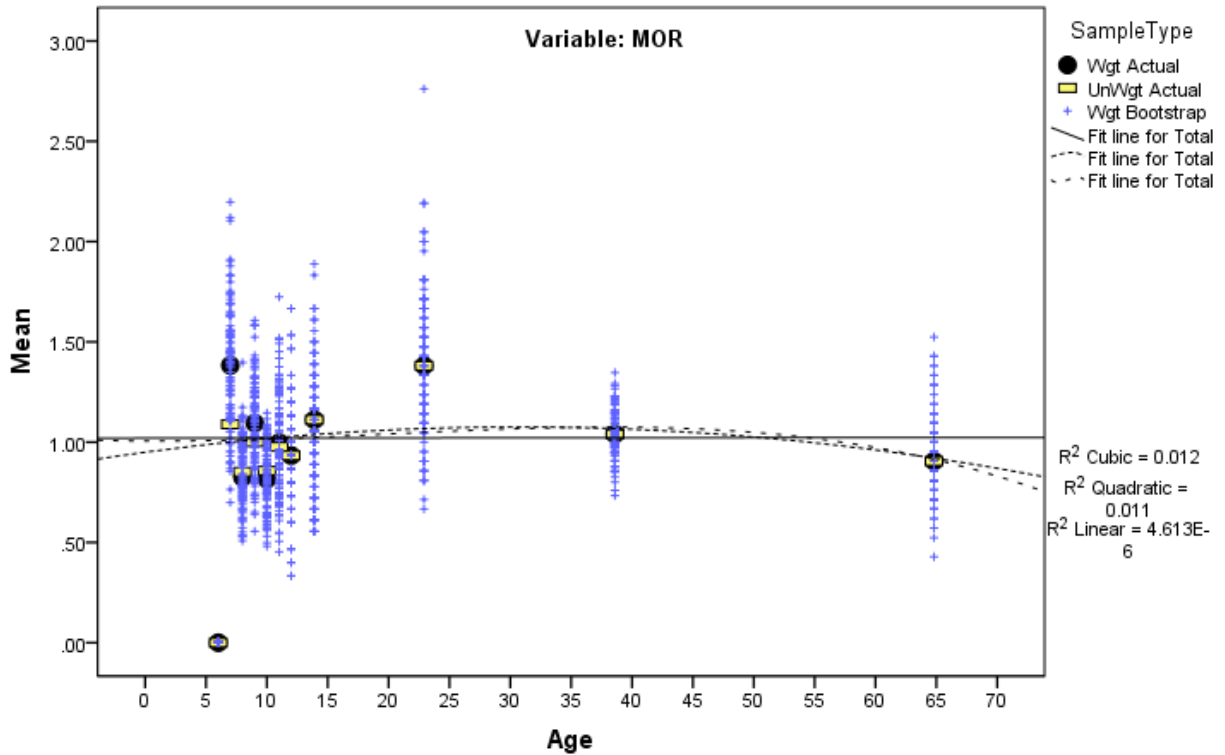
Because we had $n \geq 50$ for just ages 7, 8, 9, and 10, we expected the other samples to be affected by fairly substantial sampling error, or natural variability that causes the observed data values to depart from the true population values. To contend with this we used a statistical procedure called bootstrapping (Efron & Tibshirani, 1993; Howell, 2010) to create 100 alternative possible versions of the existing age-based data sets for each variable. These bootstrap samples show what the sample could have looked like when drawing samples of the same size from the same population. In essence, they create empirical sampling distributions that show how much the sample estimate could vary by chance alone. Although 100 bootstrap samples would be considered small in the statistical literature (with 1,000 or more samples being common), this sample size was sufficient for our purposes.

Next we fit regression equations to predict means for all the scores on the Page 1 and Page 2 profiles from age. As the dependent variable we used scores from the actual age based samples (weighted in an effort to balance contributions across countries and also unweighted) as well as the 100 bootstrap samples. In the regression models, we gave full weight (1.0) to the best balanced actual sample, a mid-weight (.75) to the unweighted actual data, and lower weight (.50) to each of the bootstrap samples. Regressions were computed for each score. And for each we modeled the linear, quadratic, and cubic functions of age (i.e., Age, Age², and Age³). After predicting the score means, we predicted the score standard deviations. Our goal was to find the best fitting function that made developmental sense. In the inferential norming literature a key step is to review the alternative regression results both to see how much prediction increases when moving from a linear to a quadratic and then a cubic function and, most importantly, to see if the various alternative regression models make developmental sense.

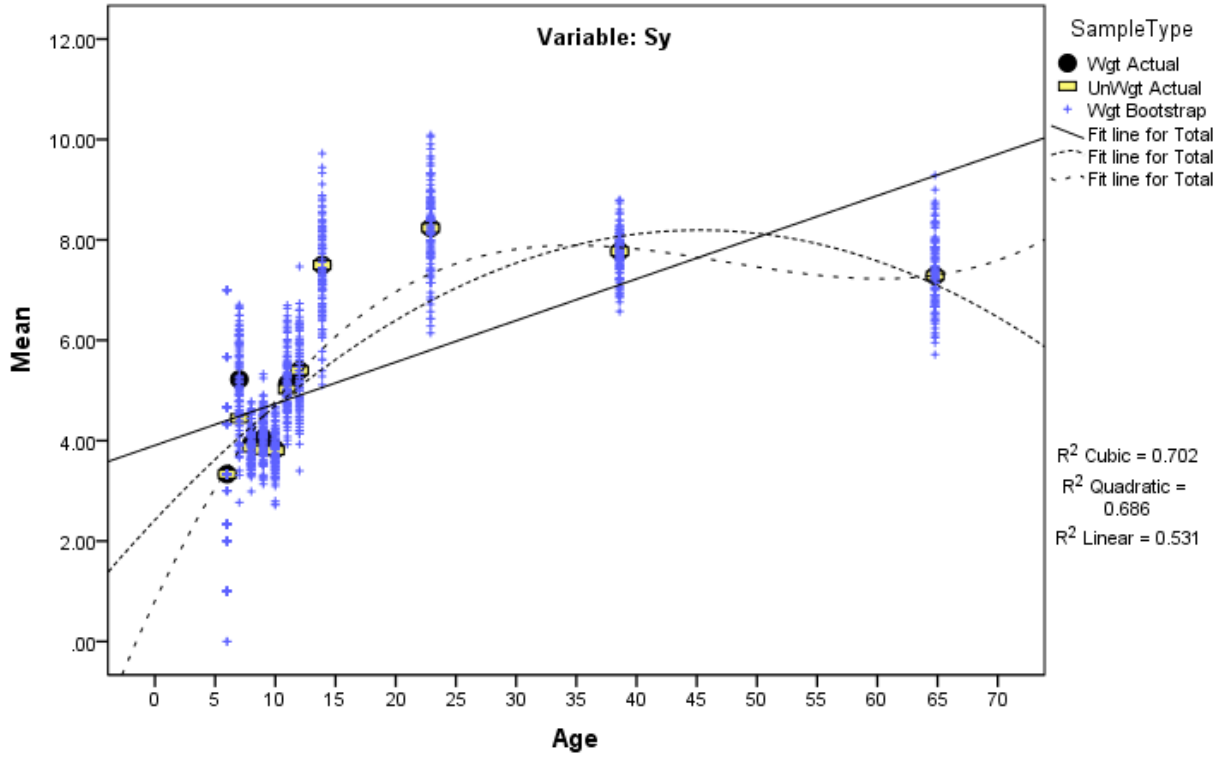
This was the procedure we followed. Using the visual scatterplot of age with the target scores as well as the statistical information about model fit, one of us (GJM) started and identified the best fitting function for every mean and standard deviation for every score. At times, the judgment was that none of the three statistical functions was optimal but that a variation of one or two made sense (e.g., an average of the quadratic and cubic function or a linear function but with a slightly higher [or lower] intercept and a flatter [or steeper] slope). Then two of us (DJV and LG) independently reviewed the initial judgments to make a determination about the function that best captured a genuine developmental trend. Then through discussion we reconciled the relatively few instances when we had a disagreement. Once we decided on the appropriate developmental function, we ran the final analyses to predict age-based means and standard deviations for every variable.

The figures below illustrate the data and our decisions. In each figure, the black circle indicates the weighted average score designed to better balance contributions across countries. The yellow rectangle is the unweighted actual sample mean (for adults, these two are equivalent). The blue crosses indicate the means from each of the 100 bootstrap samples. The solid line designates the linear regression estimate, the dotted line designates the quadratic function, and the dashed line indicates the cubic function. In the

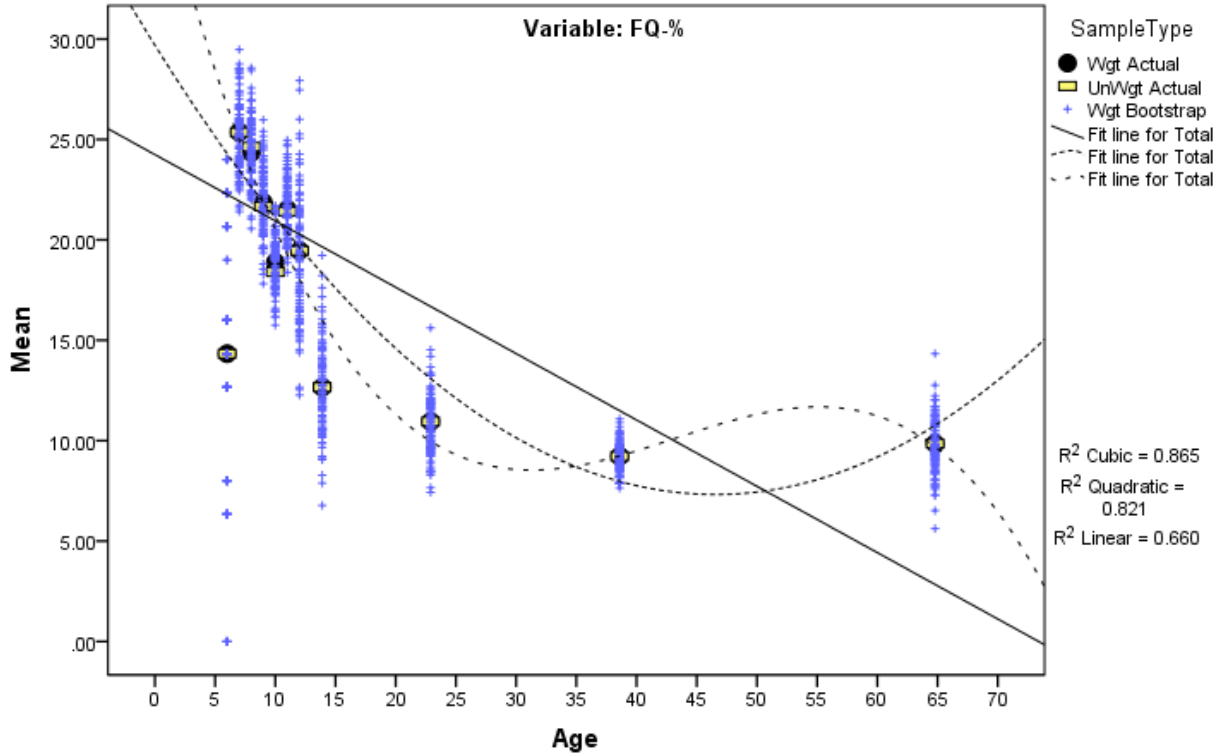
first scatterplot for the Morbid (MOR) score, all three regression lines are very close to each other. Our decision here was that the flat linear model provided the best logical fit. In the second plot for Synthesis (Sy), we concluded that the best fitting line for the child and adolescent samples was given by the quadratic function. In the third plot for Form Quality Minus Percent (FQ-%) our judgment was that the best fit was obtained by taking an average of the quadratic and cubic functions, such that the final predicted values lie halfway between each of those two curves.



Cases weighted by RegressionWeight

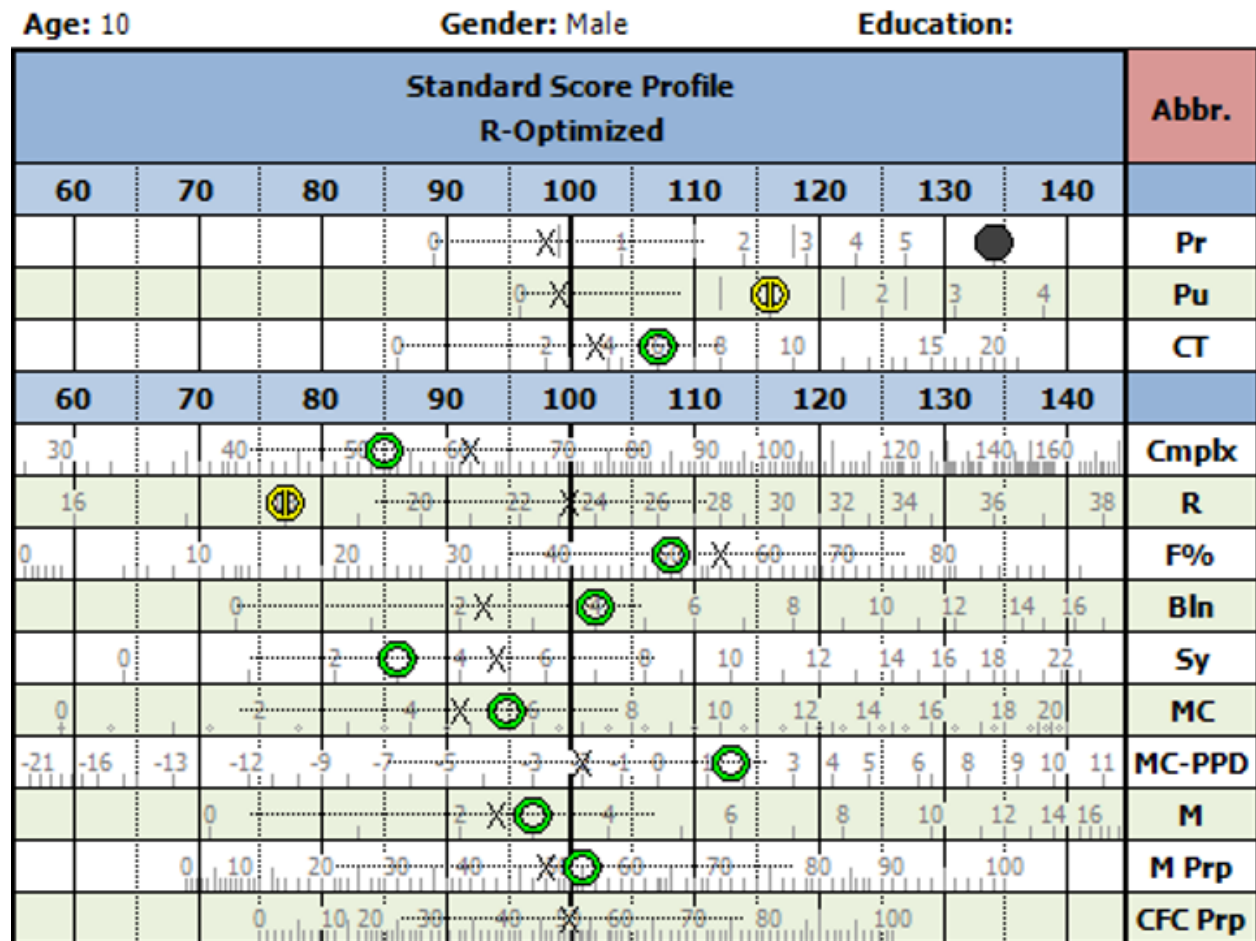


Cases weighted by RegressionWeight



Cases weighted by RegressionWeight

As a result of this process, we have provisional R-PAS norms for children and adolescents, consisting of predicted means and standard deviations for every profiled variable across ages 6 to 17 in one-year increments. When entering a protocol for scoring, if the examiner has used R-Optimized administration and the R-PAS FQ tables, there will be the option to select one of these specific ages. When the case is processed, the output provides the age-specific standard scores for each variable. For purposes of developmental comparison, the child or adolescent's standard score relative to adult expectations is provided too. The Page 1 and Page 2 Profiles plot the child or adolescent's raw score on the standard adult normative grid. However, each score is accompanied by overlays that show the expected normative range for the child's age. The overlays include an X that designates the predicted mean score and dotted line whiskers that extend one standard deviation above and below the mean. Complexity-Adjusted standard scores are also available, with an Appendix that provides normed results for all variables. An enlarged view of a portion of the Page 1 Profile output is given as an example below.



When the new normative values are combined to create the age bins that best match those that had been used by Meyer, Erdberg, and Shaffer (2007) and in our initial normative overlays, the new data are typically very similar to the old data. In instances when the two sets of values differ, the variable either was designated as "unstable" in the earlier norms, had values in the previous norms that did not show developmental trends but were consistently elevated or depressed relative to adult norms, or the new norms corrected developmental irregularities that were present in the previous norms (e.g., in the previous norms, there were developmental progressions for many determinant-related variables, including F%, Blends, CBlend, YTV, PPD, Y, V, C', and r, but the 13- to 18-year-old norms overshot the

adult norms; now with the new norms the age bands show a dimensional progression toward the adult values).

We are in the process of collecting official new reference data for children and adolescents and are interested in partnering with people who might be able to collect such data from their country or region, even if it might be just a small sample. If you are interested in exploring this possibility with us, please visit the Help with Normative Data Collection link on your R-PAS control panel for more information. We require all examiners to pass proficiency exams in administration and coding before data collection begins and we provide substantial support to help ensure this will be successful, including a free six-week training course online.

Note to Examiners with Existing Child and Adolescent Protocols in the R-PAS Scoring

Program: Child and adolescent protocols that either were obtained using CS administration guidelines or scored using the CS FQ tables still have to use the old normative overlays for ages 5 to 8, 9 to 12, or 13 to 15, as described in our Problems with Child and Adolescent Normative Data Reference Data document. These protocols have retained their existing settings in the R-PAS scoring program. However, if a protocol in the R-PAS scoring program was administered using R-Optimized administration and scored using the R-PAS FQ tables, we mechanically switched that protocol from the old norms to the new norms using existing age information entered into the scoring program. These age entries were automatically converted if they were entered as numeric values (e.g., 9, 11, 16.2) or as string variables where the age designation could be obtained from the text string (e.g., 8,2, 11:6, 15-5, 13/3). In both instances, the age entry was truncated to the year value (i.e., 7.9 became 7 not 8). However, in a small number of instances automatic conversion was not possible (e.g., when a date was entered in the age field). As a result, some child or adolescent protocols in the R-PAS scoring program are now profiled without any overlay and will need to be manually updated, which is easy to do. Simply select the protocol from the View Protocols page, click the "View selected protocol" link, click on the "Edit case information" link, choose the appropriate age in the Age Table section, and then click the "Update protocol" button at the bottom of the page. Protocols for 18-year-olds were converted to the adult norms, though users also may find it helpful to view the overlay for age 17. Protocols for 5-year-old children were not converted from the old overlays, though the current 6-year-old norms may provide a better approximation than the old norms encompassing the age range from 5 to 8 years.

References

- Efron, B. & Tibshirani, R. T. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.
- Howell, D. C. (2010). *Statistical Methods for Psychology* (8th Ed.). Belmont, CA: Wadsworth, Cengage Learning.
- Meyer, G. J., Erdberg, P., & Shaffer, T. W. (2007). Towards international normative reference data for the Comprehensive System. *Journal of Personality Assessment, 89*, S201-S216. DOI: 10.1080/00223890701629342
- Roid, G. H. (2003). *Stanford-Binet Intelligence Scales—Fifth Edition: Technical manual*. Itasca, IL: Riverside.
- Wechsler, D. (2003). *Manual for the Wechsler Intelligence Scale for Children—Fourth Edition*. San Antonio, TX: Psychological Corporation.

- Wechsler, D. (2008). *WAIS-IV: Technical and interpretive manual*. San Antonio, TX: Pearson.
- Wechsler, D. (2009). *Wechsler Individual Achievement Test—Third Edition*. San Antonio, TX: Pearson.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Rolling Meadows, IL: Riverside.
- Zachary, R. A., & Gorsuch, R. L. (1985). Continuous norming: Implications for the WAIS-R. *Journal of Clinical Psychology, 41*, 86-94. doi:10.1002/1097-4679(198501)41:1<86::AID-JCLP2270410115>3.0.CO;2-W
- Zhu, J., & Chen, H.-Y. (2011). Utility of inferential norming with smaller sample sizes. *Journal of Psychoeducational Assessment, 29*, 570-580. doi:10.1177/0734282910396323