# Woodcock-Johnson® III

## Assessment Service Bulletin Number 9

# *Woodcock-Johnson III®/Woodcock-Johnson III Normative Update* Score Differences: What the User Can Expect and Why

*Kevin S. McGrew, PhD*
*David E. H. Dailey, MS*
*Fredrick A. Schrank, PhD*

*This bulletin describes the differences between the 2000 U.S. Census projections and the 2000 U.S. Census statistics and the impact this had on the WJ III norms. It further describes how recalculation of the WJ III norms based on the final 2000 U.S. Census statistics and the use of state-of-the-art bootstrap resampling procedures resulted in the* WJ III Normative Update *(WJ III NU), a more current and accurate comparison of an individual's scores to the U.S. population. Finally, it provides recommendations for best practice use of the WJ III/WJ III NU scores in tracking individuals' performance across time, allowing users of the WJ III to have great confidence in the accuracy of the scores from this instrument.*

Riverside Publishing
A HOUGHTON MIFFLIN COMPANY

Printed in the United States of America.



**Reference Citation**

- To cite this document, use:

  McGrew, K. S., Dailey, D. E. H., & Schrank, F. A. (2007). Woodcock-Johnson III/Woodcock-Johnson III Normative Update *Score Differences: What the User Can Expect and Why* (Woodcock-Johnson III Assessment Service Bulletin No. 9). Rolling Meadows, IL: Riverside Publishing.

For technical information, please call 800.323.9540 or visit our website at www.woodcock-johnson.com.

1 2 3 4 5 6 7 8 9 10-XBS-10 09 08 07

# Overview

The *Woodcock-Johnson III Normative Update* (WJ III® NU) (Woodcock, McGrew, Schrank, & Mather, 2001, 2007) is a recalculation of the normative data for the *Woodcock-Johnson III* (WJ III) (Woodcock, McGrew, & Mather, 2001), based on the final 2000 U.S. census statistics (U.S. Census Bureau, 2005). The final 2000 census data are reflected in the norms provided by the *WJ III Normative Update Compuscore® and Profiles Program* (Compuscore) (Schrank & Woodcock, 2007) and in the documentation provided in the *WJ III Normative Update Technical Manual* (McGrew, Schrank, & Woodcock, 2007). The WJ III NU norms replace the original WJ III norms, which were based on the U.S. Census Bureau's 2000 census projections issued in 1996 (Day, 1996).

The U.S. Census Bureau's Population Projections Program issues projections of the United States resident population based on assumptions about future births, deaths, and international migration. *Census projections are estimates of the population for future dates and are subsequently replaced by census statistics*. The 2000 census statistics produced a somewhat different description of the U.S. population than was assumed from the last projections issued in 1996. For example, according to the bureau's Greg Spencer, "When we took the 2000 census, we found about 6.8 million more people than we were expecting. When we went in and looked at the sources of that growth, we found that during the late 1990s, there was more migration than we had been measuring" (Landphair, 2004, p. 1). Other unanticipated changes in the population were documented (U.S. Census Bureau, 2005), including shifts in age, sex, race, Hispanic origin, and residence. Some states grew at three times the national rate, and people tended to cluster in locations where jobs were available and climate was preferred. For users of the WJ III, the normative update (WJ III NU) provides the most current comparisons to the U.S. population.

In addition, the WJ III NU used innovative statistical advancements to calculate the new norms. The use of *bootstrap* resampling procedures (Efron & Tibshirani, 1993) allowed for estimates of uncertainty and potential bias (in the sample data) to be incorporated into the calculation of the WJ III NU norms. The bootstrap-based norm development procedures used to recalculate WJ III NU norms result in more precise estimates of an individual's tested performance.

Collectively, the refinement of the demographic characteristics in the WJ III NU norm data, based on updated U.S. census statistics, and the use of state-of-the-art statistical methods for estimating the sample statistics used to calculate norms allow users of the WJ III to have greater confidence in the accuracy of the WJ III NU-based scores.

## Purposes

Although recalculating the norms provides greater accuracy, it is important to address common questions that may arise concerning the accuracy of the methods used. The following five general categories of questions are presented and discussed.

- What differences can I expect between test scores based on the original WJ III and the WJ III NU norms? If there are score differences, are the differences more noticeable with certain age groups?
- Why are the WJ III/WJ III NU SS average differences larger for some tests at the youngest and oldest age groups?
- Did the 2000 U.S. census population demographic final statistics change enough from the year 2000 census projections to make a real difference in the scores subjects will receive on the WJ III? That is, are the WJ III NU norms needed to provide better estimates of scores for comparison to the current U.S. population?
- What is "bootstrap resampling" and, more importantly, how does this innovative statistical method provide more accurate norms than those published with the WJ III data in 2000?
- What WJ III/WJ III NU scores should be used to track and compare test performance for individuals across time? What is the suggested best practice for tracking student growth from the WJ III to the WJ III NU?

### What differences can I expect between test scores based on the original WJ III and the WJ III NU norms? If there are score differences, are the differences more noticeable with certain age groups?

To answer this question, the WJ III test scores for all 8,782 norm subjects were calculated using both the WJ III and WJ III NU scoring programs. Age-based standard scores (SS) were calculated for each test using both sets of norms. The WJ III set of standard scores for each norm subject was then subtracted from the WJ III NU standard scores (WJ III NU SS – WJ III SS = WJ III/WJ III NU SS Difference). A negative SS difference score would indicate that, for the same norm subject, his or her WJ III NU SS was lower. A positive SS difference score would indicate that, for the same norm subject, his or her WJ III NU SS was higher.

The average (median) WJ III/WJ III NU SS differences for each test for each of the 25 norm technical age groups (i.e., ages 2, 3, 4, etc.)[1] were then plotted on a graph (as a function of average chronological age), and polynomial curve fitting procedures were employed to generate smoothed curves that best fit the median WJ III/WJ III NU SS differences across age. These smoothed curves provide the optimal estimates of the WJ III/WJ III NU SS difference parameters (see WJ III NU and WJ III norm curve fitting procedures as described by McGrew, Schrank, & Woodcock, 2007 and McGrew & Woodcock, 2001).

The final average WJ III/WJ III NU SS difference estimates are reported (by chronological age) in Tables 1 and 2. Examiners can use the information presented in Tables 1 and 2 to anticipate the typical WJ III-to-WJ III NU SS differences when switching from the WJ III to WJ III NU scoring software. It must be remembered these are *average* differences due to the differences in the WJ III and WJ III NU norm calculation procedures (see McGrew, Schrank, & Woodcock, 2007). Suggestions for comparing scores for subjects originally tested and scored with the WJ III norms and subsequently retested and scored with the WJ III NU norms are presented in a subsequent section of this report. In addition, explanations for the observed differences are presented in subsequent sections of this report.

---

[1] The WJ III and WJ III NU technical age groups correspond to the age categories used to report summary statistics in the respective technical manuals.

***Table 1.***

*Average (Median) Standard Score (M = 100; SD = 15) Differences for WJ III Cognitive Test Scores (Calculated for all Norm Subjects [by Age] Based on WJ III and WJ III NU Norms)*

| | Chronological Age Groups (in years) | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Preschool (P) | | | | School-Age (S) | | | | | | | | | | | | | Adult (A) | | | | | | | | | *SD(Diff)*[a] | | |
| **Test/GIA cluster** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20–29** | **30–39** | **40–49** | **50–59** | **60–69** | **70–79** | **80–89** | **90+** | **P** | **S** | **A** |
| Verbal Comprehension | **–6** | **–4** | 0 | **2** | 1 | –1 | **–2** | **–3** | **–3** | **–3** | **–2** | **–2** | –1 | –1 | –1 | 0 | 0 | 0 | 0 | **–2** | **–4** | **–6** | **–5** | **–4** | **–3** | 0 | 3 | 7 | 5 |
| Visual-Auditory Learning | | **–4** | **–4** | **–3** | **–3** | **–3** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–3** | **–4** | **–5** | **–5** | **–4** | **–4** | 13 | 12 | 12 |
| Spatial Relations | **–9** | **–3** | –1 | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | 0 | –1 | **–3** | **–3** | **–2** | **–2** | **–3** | 4 | 11 | 9 |
| Sound Blending | | –1 | –1 | –1 | 0 | 0 | 1 | 1 | **2** | **2** | **2** | **2** | 1 | 0 | –1 | **–2** | **–2** | **–2** | **–2** | –1 | 0 | 0 | 0 | –1 | –1 | **–2** | 3 | 2 | 3 |
| Concept Formation | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | –1 | –1 | –1 | **–2** | **–2** | **–2** | –1 | 0 | 0 | 0 | –1 | –1 | **–2** | **–2** | **–2** | **–2** | **–2** | 2 | 6 | 3 |
| Visual Matching | | 0 | 0 | **–2** | **–3** | **–3** | **–2** | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | –1 | –1 | **–2** | **–2** | **–3** | **–3** | **–2** | **–2** | **–2** | **–2** | 3 | 1 | 1 |
| Numbers Reversed | | | | **–2** | 0 | 0 | 0 | –1 | **–2** | **–2** | **–2** | **–2** | **–2** | –1 | 0 | 1 | **2** | **2** | 1 | –1 | **–2** | **–2** | **–2** | **–2** | **–3** | **–3** | 8 | 8 | 8 |
| **GIA–Standard** | | | | **–3** | **–3** | **–2** | **–2** | –1 | –1 | –1 | –1 | **–2** | **–3** | **–3** | **–3** | **–3** | **–3** | **–2** | **–2** | **–3** | **–3** | **–3** | **–3** | **–3** | **–3** | **–3** | 2 | 2 | 4 |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Incomplete Words | **–10** | **–3** | **–3** | –2 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –2 | **–3** | **–4** | **–4** | **–3** | –2 | –2 | –2 | **–3** | 4 | 3 | 3 |
| Auditory Working Memory | | | **–2** | 0 | –1 | **–2** | **–2** | **–2** | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **–2** | **–3** | **–3** | **–2** | –1 | 0 | –1 | **–3** | 2 | 12 | 11 |
| General Information | **–5** | **–4** | 1 | **2** | **2** | **2** | **2** | **2** | **2** | 0 | –1 | **–2** | **–2** | **–3** | **–3** | **–3** | **–3** | **–2** | 0 | 0 | **–3** | **–5** | **–5** | **–5** | **–5** | **–4** | 2 | 7 | 5 |
| Retrieval Fluency | **–3** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | –1 | –1 | **–2** | **–2** | **–3** | **–3** | **–3** | **–3** | **–3** | **–3** | **–2** | **–2** | **–4** | **–4** | **–3** | **–2** | **–2** | 4 | 11 | 9 |
| Picture Recognition | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | **–2** | **–2** | **–2** | –1 | –1 | **–3** | **–3** | **–3** | –1 | –1 | 2 | 13 | 12 |
| Auditory Attention | **–6** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | **–2** | **–3** | **–3** | **–2** | **–2** | **–2** | **–2** | 5 | 15 | 13 |
| Analysis-Synthesis | | | **–4** | **–3** | **–2** | –1 | –1 | –1 | **–2** | **–2** | **–2** | –1 | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | –1 | –1 | –1 | –1 | 0 | 0 | 7 | 10 | 9 |
| Decision Speed | | **–6** | **–2** | 0 | –1 | –1 | –1 | –1 | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | –1 | **–3** | **–3** | **–2** | –1 | 0 | 0 | **–2** | 9 | 10 | 9 |
| Memory for Words | **–2** | **–2** | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | **–2** | **–3** | **–3** | 1 | 7 | 3 |
| **GIA–Extended** | | | **–2** | **–2** | | | | | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–3** | **–4** | **–4** | **–3** | **–2** | **–2** | **–2** | 2 | 3 | 4 |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Rapid Picture Naming | –1 | –1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | –1 | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | 11 | 14 | 12 |
| Planning | | **–4** | **–2** | 0 | 0 | –1 | –1 | **–2** | **–2** | **–2** | **–2** | **–2** | –1 | –1 | –1 | –1 | –1 | 0 | 0 | 0 | –1 | **–3** | **–4** | **–5** | **–5** | **–5** | 7 | 16 | 15 |
| Pair Cancellation | | | | | –1 | –1 | –1 | –1 | –1 | –1 | 0 | 0 | 0 | 0 | –1 | –1 | –1 | –1 | –1 | **–3** | **–3** | **–3** | –1 | –1 | 0 | 0 | | 19 | 17 |
| Memory for Names | **–2** | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **–2** | **–2** | **–2** | –1 | 0 | –1 | –1 | 8 | 13 | 13 |
| Visual Closure | **–2** | **–2** | –1 | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **–2** | **–3** | **–3** | **–3** | **–3** | **–3** | **–3** | 8 | 8 | 4 |
| Sound Patterns–Voice | | –1 | 0 | 0 | –1 | –1 | –1 | –1 | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | –1 | 0 | 1 | 1 | 1 | 2 |
| Number Series | | | | **–3** | **–2** | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | 0 | 0 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | 4 | 2 | 3 |
| Number Matrices | | | | | | | –2 | **–2** | **–2** | **–2** | **–2** | –1 | –1 | –1 | 0 | 0 | 1 | 1 | **2** | **2** | **2** | 1 | –1 | **–4** | **–5** | **–6** | | 1 | 2 |
| Cross Out | | | –1 | –1 | **–2** | **–2** | –1 | –1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | –1 | –1 | –1 | –1 | –1 | 1 | 1 | 15 | 10 | 8 |
| Memory for Sentences | –1 | **–2** | **–2** | –1 | 0 | 0 | 0 | –1 | **–2** | **–2** | **–2** | **–2** | **–2** | **–2** | –1 | –1 | 0 | 0 | 1 | 1 | 1 | 1 | –1 | **–2** | **–2** | **–2** | 2 | 1 | 2 |
| Block Rotation | | **–4** | **–2** | 0 | 0 | 0 | 0 | 0 | 0 | –1 | –1 | **–2** | **–2** | **–2** | **–2** | **–2** | –1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 | 2 | 5 |
| Sound Patterns–Music | | **–4** | **–3** | –1 | 0 | 0 | 0 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | –1 | 0 | 0 | 0 | –1 | –1 | –1 | –1 | 0 | 0 | –1 | **–2** | 2 | 1 | 1 |

*Note.* Negative difference indicates that average WJ III NU standard score is lower. Positive value indicates that average WJ III NU standard score is higher.
Bold values designate average differences that are operationally defined as significant ( <–2 or >+2).
[a] *SD(Diff)* = Standard deviation of WJ III NU SS differences for broad Preschool (P), School-Age (S), and Adult (A) age ranges.

***Table 2.***

*Average (Median) Standard Score (M = 100; SD = 15)*
*Differences for WJ III Achievement Test Scores (Calculated for All Norm Subjects [by Age] Based on WJ III and WJ III NU Norms)*

| | Chronological Age Groups (in years) | | | | | | | | | | | | | | | | | | | | | | | | | | SD(Diff)[a] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Preschool (P) | | | | School-Age (S) | | | | | | | | | | | | | Adult (A) | | | | | | | | | P | S | A |
| **Test** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20–29** | **30–39** | **40–49** | **50–59** | **60–69** | **70–79** | **80–89** | **90+** | | | |
| Letter-Word Identification | **−7** | **−5** | **−4** | **−4** | **−3** | 0 | 0 | 0 | −1 | **−2** | **−2** | **−2** | **−2** | −1 | −1 | 0 | 0 | 1 | 0 | **−2** | **−4** | **−4** | **−3** | **−2** | **−2** | **−2** | 11 | 5 | 4 |
| Reading Fluency | | | | | −1 | −1 | −1 | 1 | 0 | −1 | −1 | −1 | −1 | −1 | 0 | 0 | 1 | **2** | **2** | −1 | **−2** | −1 | −1 | **−2** | **−4** | **−4** | | 11 | 8 |
| Story Recall | −1 | **−2** | **−3** | **−3** | **−2** | **−2** | −1 | −1 | −1 | −1 | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **−2** | **−3** | **−3** | **−3** | **−3** | **−3** | 2 | 1 | 2 |
| Understanding Directions | 1 | 1 | 0 | 0 | 0 | −1 | −1 | **−2** | **−2** | **−2** | **−2** | **−2** | **−2** | −1 | 0 | 1 | 1 | 1 | 1 | −1 | **−2** | **−3** | **−3** | **−3** | **−3** | **−3** | 11 | 14 | 11 |
| Calculation | | | | **−5** | **−2** | **−2** | **−2** | **−2** | **−3** | **−3** | **−3** | **−3** | **−2** | 0 | 1 | **2** | **2** | **2** | 0 | **−3** | **−3** | **−3** | **−2** | **−2** | **−2** | **−2** | 7 | 2 | 3 |
| Math Fluency | | | | **−5** | −1 | −1 | **−2** | **−2** | −1 | −1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | −1 | **−2** | **−2** | **−2** | **−2** | **−3** | 7 | 6 | 5 |
| Spelling | | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | **−2** | | **2** | **2** | −1 | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | 2 | 5 | 4 |
| Writing Fluency | | | | **4** | **3** | 0 | −1 | −1 | −1 | −1 | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **−2** | **−2** | **−2** | −1 | −1 | −1 | | 8 | 6 |
| Passage Comprehension | | | **−10** | **−5** | **−2** | **−2** | **−2** | **−2** | **−2** | **−2** | **−2** | −1 | −1 | −1 | 0 | 0 | 0 | 0 | 0 | −1 | **−2** | **−2** | **−3** | **−3** | **−2** | **−2** | 5 | 5 | 4 |
| Applied Problems | **−3** | 0 | 1 | 1 | 1 | **−2** | **−3** | **−4** | **−4** | **−3** | **−2** | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | **−3** | **−5** | **−5** | **−5** | **−5** | **−5** | 3 | 4 | 4 |
| Writing Samples | | | | **−6** | **−4** | **−3** | **−2** | **−2** | **−2** | −1 | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **−3** | **−3** | **−3** | **−3** | **−3** | **−4** | **−4** | 9 | 7 | 7 |
| Word Attack | | **−5** | **−4** | **−3** | **−3** | **−3** | **−2** | **−2** | **−2** | **−2** | **−2** | **−2** | −1 | 0 | 1 | 1 | 1 | 1 | 0 | −1 | **−2** | **−2** | **−2** | **−2** | **−2** | **−3** | 9 | 6 | 7 |
| Picture Vocabulary | −5 | **−2** | **−2** | **−2** | **−2** | **−3** | **−3** | **−3** | **−3** | **−3** | **−2** | **−2** | **−2** | **−2** | **−2** | **−2** | **−2** | **−2** | −1 | −1 | −1 | −1 | −1 | **−2** | **−2** | **−3** | 2 | 1 | 2 |
| Oral Comprehension | **−7** | **−6** | **−3** | **−2** | **−2** | **−2** | **−2** | **−2** | **−2** | −1 | 0 | 1 | **2** | **2** | **2** | **2** | **2** | **2** | −1 | **−2** | **−3** | **−3** | **−3** | **−3** | **−3** | **−3** | 2 | 6 | 7 |
| Editing | | | | | **4** | **4** | 0 | −1 | −1 | −1 | −1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | **−3** | **−3** | **−3** | **−3** | **−2** | **−2** | **−2** | | 10 | 9 |
| Reading Vocabulary | | | | **−5** | **−2** | **−2** | **−2** | **−2** | −1 | −1 | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | **−2** | **−2** | **−2** | **−2** | −1 | 1 | 10 | 1 | 2 |
| Quantitative Concepts | | **−2** | **−2** | **−3** | **−3** | **−3** | **−3** | **−3** | **−4** | **−3** | **−3** | **−3** | **−3** | **−2** | **−2** | −1 | −1 | 1 | 1 | −1 | **−4** | **−4** | **−4** | **−3** | **−3** | **−3** | 10 | 6 | 4 |
| Academic Knowledge | **−2** | **−2** | 0 | 1 | 1 | 1 | 1 | 0 | 0 | −1 | **−2** | **−2** | **−2** | **−2** | **−2** | **−2** | 0 | 1 | 1 | 1 | 1 | 0 | 0 | −1 | **−3** | **−3** | 1 | 4 | 4 |
| Spelling of Sounds | | | | **−5** | **−2** | −1 | 0 | −1 | **−2** | **−2** | **−2** | **−2** | **−2** | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | −1 | −1 | −1 | −1 | −1 | 5 | 11 | 8 |
| Sound Awareness | | | **−5** | **−4** | **−2** | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | −1 | **−3** | **−4** | **−4** | **−5** | **−5** | **−5** | 5 | 12 | 9 |

*Note.* Negative difference indicates that average WJ III NU standard score is lower. Positive value indicates that average WJ III NU standard score is higher.
Bold values designate average differences that are operationally defined as significant ( <–2 or >+2).
[a]*SD(Diff)* = Standard deviation of WJ III NU SS differences for broad Preschool (P), School-Age (S), and Adult (A) age ranges.

# Average Differences: General Findings and Trends

In general, a review of the information in Tables 1 and 2 indicates that when a significant average WJ III/WJ III NU SS difference is present, typically the difference is that the WJ III NU SS is lower. A review of both tables also reveals a trend for the WJ III/WJ III NU SS differences to be largest at both ends since the WJ III NU SS scores are typically lower at the youngest and oldest age groups, particularly the former.

## School-Age Trends

A review of the results for the school-age population (ages 6 through 18) reveals the following general conclusions.

- When compared to the preschool (ages 2 through 5) and adult (19 years and above) age ranges, the magnitude of WJ III/WJ III NU score differences are the least different in the school-age population (ages 6 through 18).
- Thirteen tests showed few significant SS differences greater than –1, 0, +1. These included:
  - Achievement — Reading Fluency, Writing Fluency
  - Cognitive — Spatial Relations, Incomplete Words, Picture Recognition, Decision Speed, Memory for Words, Pair Cancellation, Memory for Names, Visual Closure, Sound Patterns–Voice, Number Series, Sound Patterns–Music.
- For the tests not mentioned above, with a few exceptions at select ages (–4 average SS difference at ages 9 and 10 for Applied Problems and age 10 for Quantitative Concepts), average SS differences of –2 to –3 points (with WJ III NU SSs being lower) would appear to be the general rule of thumb for the school-age population.
- Four tests in the school-age range displayed average positive SS differences (+2 SS in favor of higher WJ III NU scores) but only at select age ranges. These tests include Oral Comprehension (ages 14 through 18), Calculation (age 17 to 18), Sound Blending (ages 10 through 13), and General Information (ages 6 through 10). Slightly higher average scores (+3 or +4 SS points) are noted for two tests but only at two ages (Writing Fluency at age 6; Editing at ages 6 and 7).

## Preschool and Adult Age Group Trends

A review of the results in Tables 1 and 2 for ages 2 to 5 (Preschool) and 19 and above (Adult) indicates that the most noticeable WJ III/WJ III NU SS differences (greater than or equal to –4 SS points) are to be expected at the youngest and oldest age groups for some, but not all, tests. Also, in general, the differences become larger as the subjects become younger (e.g., Letter-Word Identification average SS differences are –4 for ages 4 to 5, –5 at age 3, and –7 at age 2).

Visual-Auditory Learning is a good example. For this test, at ages 3 to 4, the WJ III NU SS is, on average, 4 SS points lower than the WJ III SS. From ages 5 to 50, the average difference is most frequently –2 SS points. After age 50, the SS differences again increase to the range of –4 to –5. Examiners should use the information in Tables 1 and 2 to acquaint themselves with anticipated typical SS differences when moving from the WJ III to WJ III NU norms, paying particular attention to certain tests at the youngest and oldest age ranges.

**Why are the WJ III/WJ III NU SS average differences larger for some tests at the youngest and oldest age groups?**

The answer to this question is presented in subsequent sections of this report. Briefly, in addition to changes in the U.S. population demographics derived from the 2000 census projections and those based on the final 2000 census statistics (U.S. Census Bureau, 2005), the most significant average SS differences are the result of the application of recent advances in the calculation of norms (bootstrap sampling methods), particularly at the extreme age groups. These innovative and advanced statistical procedures allow for the development of norm curves and tables that are more accurate and precise, particularly at the beginning and end of each test's age-related stream of data—areas where norm calculation procedures typically have been most difficult for test developers.[2] Detailed explanations follow in this report.

## Range of Average Differences: General Findings and Trends

At the far right end of Tables 1 and 2 are the average standard deviations (*SD*) of the obtained WJ III/WJ III NU SS differences ($SD_{Diff}$) for the three broad age-based norm groups (P = Preschool, S = School-Age, A = Adult). For each test, the WJ III/WJ III NU SS $SD_{Diff}$ values designate the range within which 68% of the WJ III/WJ III NU SS differences occur. The results for Incomplete Words are used here to demonstrate how to use and interpret this information.

As reported in Table 1, the $SD_{Diff}$ for Incomplete Words is 4 SS points (Preschool), 3 SS points (School-Age), and 3 SS points (Adult), respectively. When combined with the average SS differences reported for each chronological age (Table 1), the following illustrative interpretations are appropriate. In each example, the average WJ III/WJ III NU SS difference indicates that the WJ III NU Incomplete Words scores are lower than the WJ III Incomplete Words scores.

### Age 5

The expected average SS difference at age 5 is –2 SS points, with 68% of the difference scores being within a range of 4 SS points. This produces an expected range of scores from –6 SS (–2 minus 4) to +2 SS points (–2 plus 4).

### Age 12

The expected average SS difference at age 12 is –1 SS point. The range within which 68% of the difference scores occur is 3 SS points. The proper interpretation is that the Incomplete Words WJ III NU SSs at age 12, on average, are expected to be –1 SS points, and 68% of the scores will range from –4 SS (–1 minus 3) to +2 SS points (–1 plus 3).

### Age 45

The expected average SS difference at age 45 is –4 SS points. The range within which 68% of the difference scores are found is 3 SS points. Thus, the proper interpretation of Incomplete Words at age 45 is that, on average, WJ III/WJ III NU SS differences are expected to be –4 SS points but can range from –7 SS (–4 minus 3) to –1 SS points (–4 plus 3).

---

[2] The explanation of the bootstrap norm calculation procedure is explained in greater detail later.

A review of the $SD_{Diff}$ information in Tables 1 and 2 indicates that no single rule of thumb can be applied across all WJ III/WJ III NU test score comparisons. The 68% range of typical WJ III/WJ III NU SS difference scores varies as a function of tests and age groups and, for some tests, varies across age groups for the same test. For example, Sound Blending has a very consistent $SD_{Diff}$ value across all ages (2 to 3 SS points). In contrast, the 68% range of difference scores for Retrieval Fluency can vary from 4 to 11 SS points. Users should consult the complete set of information in Tables 1 and 2 to ascertain the expected average WJ III/WJ III NU SS differences and general range of differences typical for each test.

**Did the 2000 U.S. census population demographic final statistics change enough from the year 2000 census projections to make a real difference in the scores subjects will receive on the WJ III? That is, are the NU norms needed to provide better estimates of scores for comparison to the current U.S. population?**

The answer to this question is "yes." Figure 1 presents a "broad stroke" visual answer to this question, using the largest norm bases (age norms) for analysis. To obtain this figure, the final age-norm subject weights used for the WJ III and WJ III NU were subtracted (WJ III age-norm weight – WJ III NU age-norm weight = WJ III/WJ III NU age-norm weight difference) for each norm subject. The histogram in Figure 1 is a plot of the WJ III/WJ III NU norm weight differences. Negative values designate subjects whose WJ III NU weights increased from their WJ III age-norm weighting (they were made to count *more* in the calculation of the WJ III NU age norms). Positive values designate subjects whose WJ III NU weights decreased from their WJ III age-norm weighting (they were made to count *less* in the calculation in the WJ III NU age norms).

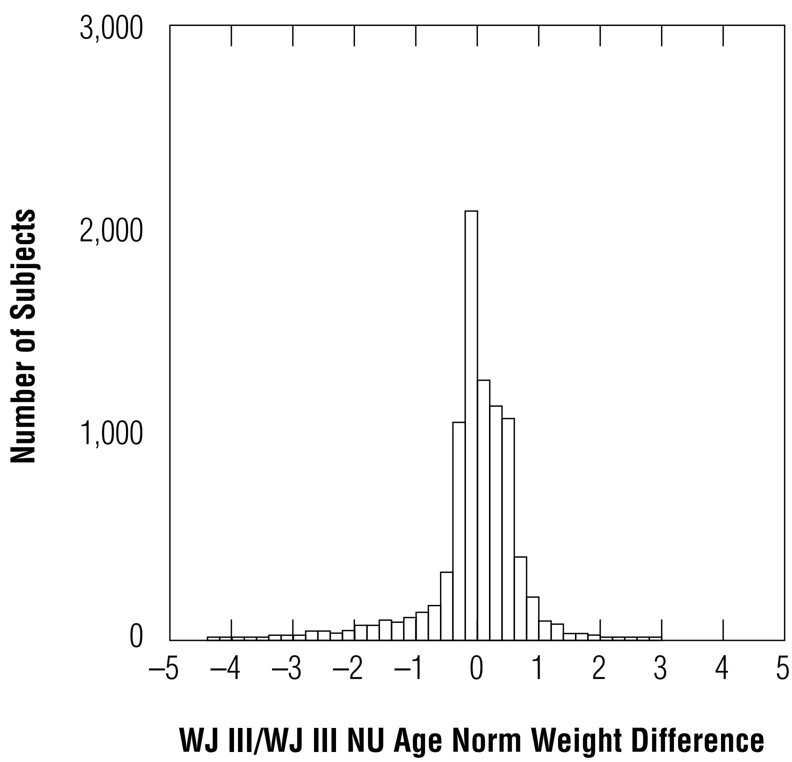



***Figure 1.***
*Distribution of WJ III/WJ III NU subject weight differences.*

In simple terms, if no differences were present between the year 2000 U.S. population demographic projections and statistics, the plot in Figure 1 would be characterized by all subject weights being packed around a value of zero. A review of the distribution in Figure 1 indicates that the majority of subjects did not change appreciably in their contribution to the calculation of WJ III and WJ III NU norms. However, there is a spread of WJ III/WJ III NU age-norm subject weight differences that suggests that when a norm subject's weighted contribution to the norms is summarized, vis-à-vis his or her final age-norm weight, a significant number of subjects' weighted contributions to the norm data changed. The spread of weight differences in Figure 1 indicates, at a broad-stroke level, the answer to the question asked in this section is "yes."

The following sections provide more in-depth insights into the nature of differences between the year 2000 U.S. census projections and final census statistics. More importantly, these projections-versus-statistics differences resulted in norm subject weight changes that indicate that the WJ III NU norms more accurately represent the performance of individuals tested today.

# Description of Year 2000 U.S. Population Projection-Versus-Statistics Changes and the Impact on the WJ III and WJ III NU Subject Weights

Table 3 summarizes the U.S. census projections/statistics changes for the adult portion of the WJ III norm sample.[3] A review of Table 3 reveals noticeable differences in the U.S. adult population in the sampling domains of Community Size, Hispanic, Education, Occupational Status, and Occupation. For illustrative purposes, detailed descriptive summary statistics are presented here only for the adult portion of the WJ III/WJ III NU norm sample. Due to space limitations, all noticeable changes in all norm bases (Preschool, Grades K through 12, College/University, Adult) cannot be described here. However, select important trends are described. More importantly, the total impact of the WJ III/WJ III NU subject weight changes in each norm basis group is summarized.

As seen in Table 3, U.S. Community Size category changes (due to changes in how the U.S. census reported the categories) of more than 2% were noticed in all three categories which, in turn, resulted in noticeable changes[4] in the community size subject weights applied to subjects classified as living in urban clusters (0.524/1.350 = .39 proportional weight change) and rural settings (1.297/.747 = 1.73 proportional weight change).[5] Given the change in the U.S. census system, significant changes in community size subject weights (greater than or equal to 20% proportional weight changes) were also noted in the urban cluster and rural category weights in the preschool and grades K through 12 norm samples.

The percentage of the U.S. adult population classified as Hispanic increased 2.5% from the year 2000 census projections to the year 2000 census final statistics, an increase resulting in a proportional weight change of 1.29 for all adult Hispanic subjects in the calculation of the WJ III NU norms (see Table 3). In other words, Hispanic subjects' scores counted 29% more in the calculation of the adult WJ III NU norms when compared to their contribution to the adult WJ III norms. A notable increase in the U.S. population classified as Hispanic in the final census statistics also occurred in all other norm bases groups. Increases in the proportion of the U.S. population classified as Hispanic were +5.4% (preschool), +3.8% (grades K to 12), and +1.5% (college/university). The corresponding Hispanic weights increased proportionately by a magnitude of 1.33, 1.34, and 1.19 respectively for these three norm groups. Clearly, one of the most noticeable changes between the year 2000 U.S. population projections and the year 2000 U.S. population final statistics, which is now incorporated in the WJ III NU norms through a change in Hispanic subject weights, is the increased proportion of U.S. individuals classified as Hispanic. The scores for Hispanic subjects are now weighted more (by approximately 20% to 30% for each Hispanic subject) in the calculation of the WJ III NU norms.

[3] Complete adult population table information can be found on page 22 (Table 2-7) of the WJ III Technical Manual and page 26 (Table 2-8) of the WJ III NU Technical Manual.

[4] Significant changes in subject weights are operationally defined as a proportional change in the weight of a magnitude of 20% or more (see Table 3). Thus, weight changes greater than or equal to 1.20 or less than or equal to .80 are highlighted in the final column of Table 3.

[5] Changes in subject weights are a function of changes in the U.S. census projections and statistics *and* the composition of the norm data. Thus, significant percentage changes in U.S. census figures did not always translate to large changes in WJ III-to-WJ III NU subject weights. Conversely, relatively small changes in the U.S. census projections/statistics might produce larger subject weight changes due to the composition of the norm data. In most cases, noticeable U.S. census percentage changes did result in corresponding significant changes in subject weights.

***Table 3.***
*Changes in U.S. Adult Year 2000 U.S. Census Projections (WJ III) and Estimates (WJ III NU)*

| | WJ III | WJ III NU | | WJ III | WJ III NU | |
|---|---|---|---|---|---|---|
| **Sampling Variable** | **Percent in U.S. Population** | **Percent in U.S. Population** | **Percentage Difference[a]** | **Subject Weight** | **Subject Weight** | **Proportion Weight Change[b]** |
| **Census Region** | | | | | | |
| Northeast | 19.0 | 18.9 | | 0.771 | 0.766 | |
| Midwest | 23.1 | 22.5 | | 1.068 | 1.038 | |
| South | 35.5 | 36.0 | | 1.100 | 1.144 | |
| West | 22.4 | 22.6 | | 1.043 | 1.020 | |
| **Community Size[c]** | | | | | | |
| Urbanized Area | 64.3 | 68.4 | **+ 4.1** | 1.024 | 1.078 | |
| Urban Cluster | 17.7 | 10.7 | **– 7.0** | 1.350 | 0.524 | **0.39** |
| Rural | 18.0 | 20.9 | **+ 2.9** | 0.747 | 1.297 | **1.73** |
| **Sex** | | | | | | |
| Male | 47.9 | 48.5 | | 1.136 | 1.166 | |
| Female | 52.1 | 51.5 | | 0.901 | 0.882 | |
| **Race** | | | | | | |
| White | 83.6 | 82.6 | | 0.987 | 0.968 | |
| Black | 11.7 | 12.0 | | 1.029 | 1.120 | |
| American Indian | 0.8 | 0.9 | | 0.620 | 0.675 | |
| Asian and Pacific Islander | 3.9 | 4.5 | | 1.454 | 1.694 | |
| **Hispanic** | | | | | | |
| Yes | 10.0 | 12.5 | **+ 2.5** | 1.058 | 1.368 | **1.29** |
| No | 90.0 | 87.5 | **– 2.5** | 0.994 | 0.963 | |
| **Education** | | | | | | |
| <9th Grade | 8.9 | 5.8 | **– 3.1** | 1.528 | 0.920 | **0.60** |
| <High School | 14.0 | 10.0 | **– 4.0** | 1.145 | 0.760 | **0.66** |
| High School | 32.5 | 31.7 | | 1.239 | 1.150 | |
| 1 to 3 Years of College | 24.1 | 27.3 | **– 3.2** | 0.985 | 1.234 | **1.25** |
| Bachelor's Degree | 13.7 | 16.8 | **+ 3.1** | 0.707 | 1.061 | **1.50** |
| Master's Degree or Higher | 6.5 | 8.4 | | 0.637 | 0.561 | |
| Not Available | — | — | — | — | — | — |
| **Occupational Status** | | | | | | |
| Employed | 64.2 | 62.7 | | 1.081 | 1.108 | |
| Unemployed | 4.1 | 3.4 | | 0.452 | 0.349 | **.77** |
| Not in Labor Force | 31.7 | 33.9 | **+ 2.2** | 1.006 | 1.006 | |
| Not Available | — | — | — | — | — | — |
| **Occupation** | | | | | | |
| Professional/Managerial | 28.1 | 33.8 | **+ 5.7** | 0.971 | 1.135 | |
| Technical/Sales/Administrative | 29.9 | 25.3 | **– 4.6** | 0.981 | 0.875 | |
| Service | 13.5 | 16.3 | **+ 2.8** | 0.718 | 0.883 | **1.23** |
| Farming/Forestry/Fishing | 2.9 | 1.6 | | 0.866 | 0.470 | **0.54** |
| Precision Product/Craft/Repair | 10.8 | 10.2 | | 1.310 | 1.145 | |
| Operative/Fabricator/Laborer | 14.4 | 12.8 | | 1.462 | 1.221 | |
| Not Available | — | — | — | — | — | — |
| **Foreign Born[d]** | | | | | | |
| Native | | 85.8 | | | 0.935 | |
| Foreign | | 14.2 | | | 1.719 | |
| Not Available | | — | | | — | — |

*Note.* See the WJ III and WJ III NU Technical Manuals for more complete information.
[a] Only WJ III/WJ III NU percentage differences of *+2%* or more are reported.
[b] Only WJ III/WJ III NU subject proportional weight differences of *20%* or more are reported (see text).
[c] At the time the WJ III was standardized, the U.S. census used the categories of Central City and Urban Fringe, Larger Community and Associated Rural Area, and Smaller Community and Associated Rural Area. For the WJ III NU, the WJ III categories were converted to the new U.S. census categories used in this table.
[d] Foreign Born was a new demographic added to the WJ III NU sample demographics. It was not used in the WJ III.

No significant changes are reported for the demographic Race variable in the adult portion of the WJ III norms (see Table 3). However, significant increases in the proportional contribution of American Indian (1.36 proportional weight change) and Asian and Pacific Islander (1.55 proportional weight change) subject scores occurred in the calculation of the WJ III NU college/university norms.

The U.S. adult population census projections and statistics changed (approximately 3% to 4%) in 4 of the 6 Education categories (see Table 3). These changes resulted in significant proportional weight changes in the educational subcategories of <9th grade (0.60), <high school (0.66), 1 to 3 years of college (1.25), and bachelor's degree (1.50). Similarly, a noticeable change in the occupational status is noted (+2.2% increase in adults classified as not in labor force). The change in the U.S. occupational status percentages, when combined with the composition of the WJ III norm data, resulted in a significant 0.77 proportional reduction in the weights assigned to adult subjects in the unemployed category. Changes in the census projections/statistics for the occupational composition of the U.S. adult population were noted in the categories of professional/managerial (+5.7% increase), technical/sales/administrative (–4.6%), and service (+2.8%). The shift in occupational status in the adult U.S. population resulted in significant proportional weight changes for two of the occupational status categories used in the calculation of the adult WJ III NU norms (technical/sales/administrative service proportional weight change = 1.23; farming/forestry/fishing proportional weight change = .54).

To increase the precision of the WJ III NU norm data for all norm group bases (Preschool, School-Age, University, Adult), the Foreign Born status of all subjects was included for the first time, resulting in the introduction of a new weighting statistic in the calculation of each subject's final norm weight.

Also of note was a slight increase in the percentage of Grade K to 12 students who were classified as home schooled (increased from 1.5% to 2.2%). Given that the original WJ III norm data included 1.1% home-schooled subjects, these students had their NU norm weights increased from 1.339 to 1.926 (proportional weight increase of 1.44).

The above review and discussion of changes between the year 2000 U.S. census population projections and statistics on select demographic variables (and resulting proportional changes in subject weights used in the calculation of the WJ III NU norms) indicate the WJ III NU scores will provide a more accurate comparison of individual scores to the current U.S. population. Given space limitations, it is impossible to review and discuss all possible U.S. demographic and WJ III/WJ III NU subject weight changes across all four norm bases. Thus, an answer to the primary questions given in this section can best be addressed via the global change statistics presented below.

As reported in the WJ III and WJ III NU Technical Manuals, each subject's weight (for each of the respective norm group bases) is the product of his or her individual subweights for each of the norm sampling variables. As can be seen in Table 3, each adult subject's contribution to the norm data is a product of his or her individual weights for nine different sampling variables (Census Region, Community Size, Sex, Race, Hispanic, Education, Occupational Status, Occupation, and Foreign Born). The multiplication of nine different weight values will produce, for many subjects, a noticeably different single adult subject weight, even if each of the subweights change only slightly.

To assess the magnitude of the changes between the projected and final adult subject statistical weights (as summarized in Table 3), a correlation was calculated between each WJ III adult subject weight and the recalculated WJ III NU adult subject weight. The obtained correlation was .43. If there were no major differences between the year 2000 U.S. population projections and statistics, one would expect very high correlations (with a correlation of +1.0 indicating no major population change at all).

Similar WJ III/WJ III NU subject weight correlations were calculated for the three other WJ III NU norm bases groups (preschool, grades K to 12, college/university). The WJ III/WJ III NU subject weight correlations for these three norm bases were .58, .63, and .46, respectively. The moderate to moderately high WJ III/WJ III NU subject weight correlations across all four norm bases groups (.43 to .63), which reinforce the value of updated WJ III NU norms, result from the fact that final subject weights incorporate the multiplicative cumulative impact of year 2000 U.S. census projection versus statistic changes across 8 (preschool), 9 (grades K to 12), 7 (college/university), and 9 (adult) demographic population characteristics. When the primary U.S. census demographic characteristics are mathematically combined to reflect the multivariate and mosaic nature of the U.S. population, significant changes have occurred between the year 2000 census projections (used in the WJ III) and the year 2000 census final statistics (used in the WJ III NU). The WJ III NU norms, which more accurately capture the complex makeup of the U.S. population, are reflected in the NU norms via the reweighting of all subjects to match the final year 2000 U.S. census statistics, which were made available in 2005.

## Caveats

It is important to recognize what the WJ III NU norms are not.

- No new norm/standardization data were gathered. The 8,782 subjects used in the construction of the NU were the same subjects tested during the standardization of the WJ III. No new subjects were recruited and/or added to the norm pool. The 8,782 WJ III subjects were reweighted to match changes between the year 2000 U.S. census population projections and statistics. The original WJ III norm subjects' scores were weighted to approximate the 2000 U.S. census projections. The WJ III NU norms use the same obtained subject *W* scores but apply different subject weights (as explained previously) so that each subject is represented according to the population demographics based on the year 2000 final U.S. population statistics.[6]
- With a few exceptions,[7] the original WJ III obtained *W* scores for all tests are identical for subjects in the calculation of the WJ III and WJ III NU norms. The difference is that the weights are applied to each subject's test/cluster *W* score for the WJ III and WJ III NU norms. In simple terms, how much each norm subject's scores "count" and influence the calculation of the average or typical score for any test/cluster (its norm) may have shifted from the WJ III to WJ III NU to more accurately reflect his or her relative contribution to the final year 2000 U.S. census statistics.

---

[6] However, in the course of new item validation/calibration for the *WJ III NU Tests of Achievement Form C/Brief Battery* (Woodcock, Schrank, Mather, & McGrew, 2007), 1,700+ new subjects were used to calibrate/validate 700+ new items.

[7] All speeded tests, Writing Samples, Understanding Directions, Story Recall, Sound Blending, and Editing tests. See McGrew, Schrank, & Woodcock, 2007.

- Given that no new data were gathered, the WJ III NU norms do not incorporate population changes due to sociocultural events or changes that may have occurred since the WJ III was published. For example, any changes in the school-age population due to recent educational reform initiatives (e.g., NCLB or Reading First) or changes in the adult population (e.g., better health care) would not be reflected in the WJ III NU norms. Only collecting completely new norm data can capture such possible shifts in the population. The NU should be considered a demographic update.

**What is "bootstrap sampling" and, more importantly, how does this innovative statistical method provide more accurate norms than those published with the WJ III data in 2000?**

As described in the WJ III and WJ III NU Technical Manuals (McGrew & Woodcock, 2001; McGrew, Schrank, & Woodcock, 2007), the development of test norms requires the establishment of the normative (average) score for each measure for subjects at each specific age (age norms) or grade (grade and university norms) where normative interpretations are intended. In the WJ family of instruments, this normative score is called the reference *W* score (REF *W*). When plotted as a function of chronological age (or grade), the REF *W* scores assume the characteristic of developmental growth curves. These test and cluster REF *W* curves are visual-graphic representations of the average performance of subjects at every age (or grade) for the effective use of the specific measure.

The REF *W* curves serve as the foundation for the age/grade equivalent, relative performance index (RPI), and instructional range interpretation features in the WJ III and WJ III NU. In addition, when the standard deviations (*SD*) of the scores at each age are plotted as a function of age/grade, the resultant curves represent the *SD* values that, when combined with the REF *W* values, provide the foundation for the calculation of all other "relative standing in a group" score metrics (e.g., standard scores and percentile ranks [PR]).

## Construction of the WJ III Norms: Letter-Word Identification Example

As described in the WJ III Technical Manual (McGrew & Woodcock, 2001), REF *W* values for a given measure are obtained from smoothed curves that pass through sample-based data points that each represent the average REF *W* values of successively ordered (by age or grade) groups or blocks of 50 norm-sample subjects. The description of this step, as well as all subsequent steps described in this section, is aided by inspecting Figures 3 and 4 (WJ III Letter-Word Identification test example). The traditional (nonbootstrap) process used in the calculation of the WJ III Letter-Word Identification age-based norms is used here to demonstrate the benefits of the improved, state-of-the-art, bootstrap-based procedures employed in the calculation of the WJ III NU norms.

Step 1 is the sorting by chronological age of all 8,648 norm subjects with Letter-Word Identification scores, represented by the cylinder, and is portrayed in Figure 2 vis-à-vis the visual line representation where subjects are arranged from the youngest (far left of dashed line) to the oldest (far right of dashed line).

Step 2 is the division of the sorted set of norm subjects into successive groups or blocks of 50 subjects. In Figure 2, each block is visually represented by the area between adjacent vertical hash mark lines. Thus, as Figure 2 portrays, the first demarcated set of the youngest subjects is at the far right of the dashed line of 8,648 age-sorted subjects. The end result of Step 2 is a set of blocks or groups of 50 subjects successively ordered from the youngest to oldest age blocks.

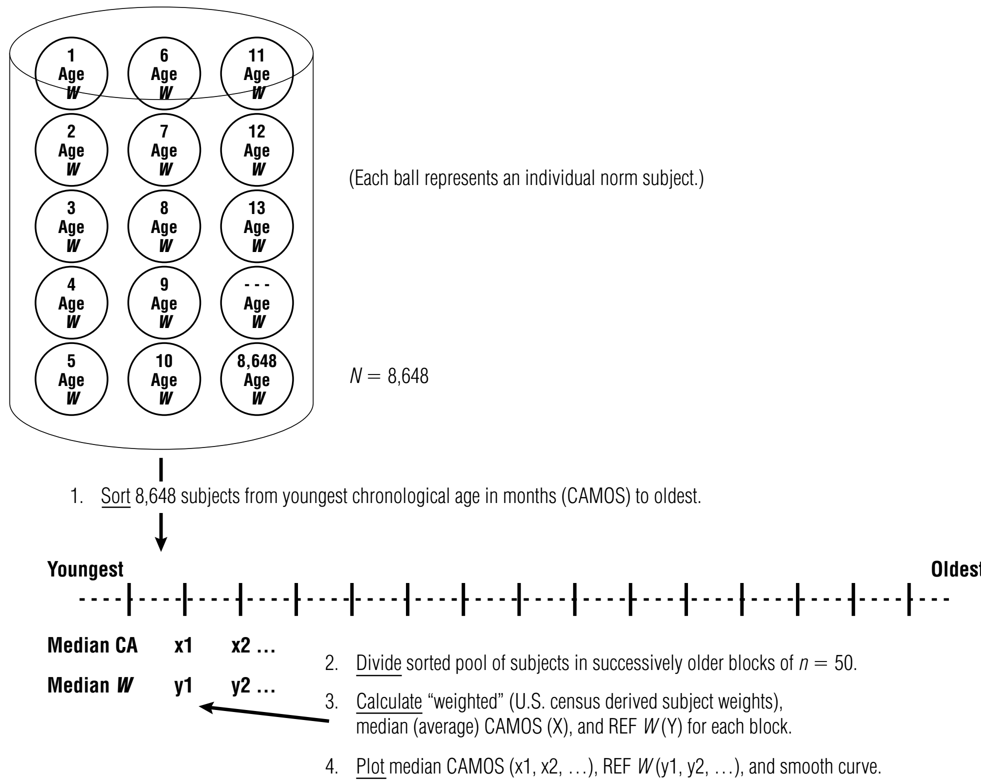



***Figure 2.***
*WJ III (traditional) norm calculation procedures example for Letter-Word Identification—Steps 1–4.*

Step 3 is the calculation of the weighted (using each subject's 2000 U.S. census weight employed during the WJ III) chronological age (in months) and Letter-Word Identification *W* score for each block (see Step 3 in Figure 2). Why and how are the weighted chronological age and *W* score values for each successive block of 50 WJ III subjects obtained?

As described in the WJ III Technical Manual, subjects in the norm sample are assumed to be randomly sampled from the population across the country. During the WJ III national standardization, efforts were made to ensure that the demographic characteristics of the norm sample matched those of the nation as a whole, given the practical constraints of applied large-scale sampling procedures (McGrew & Woodcock 2001). However, sampling efforts are rarely successful in perfectly capturing a nationally representative sample. Unless the entire population is surveyed, every norm sample is always an inexact representation of the population against which test scores will be compared (this is commonly referred to as sampling error) (Woodcock, 1994). The amount of sampling error is often unknown, and therefore, the certainty of the plotted data points is unknown. The uncertainty of the sample-based estimates is a major source of potential imprecision in the norms in most norm-referenced tests. The new bootstrap-based procedures employed in the calculation of the WJ III NU norms, which are described in subsequent sections of this document, specifically address this issue by providing estimates of the degree of certainty in the sample-based statistics.

To address the realities of sampling procedures that are less than 100% perfect, test developers traditionally statistically weight each subject's scores to represent the cumulative effect of the subject's over- or under-representation (relative to the population) within the norm sample, along several demographic characteristics. The WJ III subject norm weights for the demographic variables are reported in Tables 2-4 through 2-7 of the WJ III Technical Manual (McGrew & Woodcock 2001). As described in the WJ III Technical Manual, the census-weighted average (median) chronological age and REF *W* scores are calculated for each block of successively age- (or grade-) based sorted 50 subjects. These pairs of age/*W* score values for all blocks serve as the raw material for Step 4 (see Figure 2), the plotting and calculation of the WJ III norm REF *W* curve for Letter-Word Identification (see Figure 3).[8]

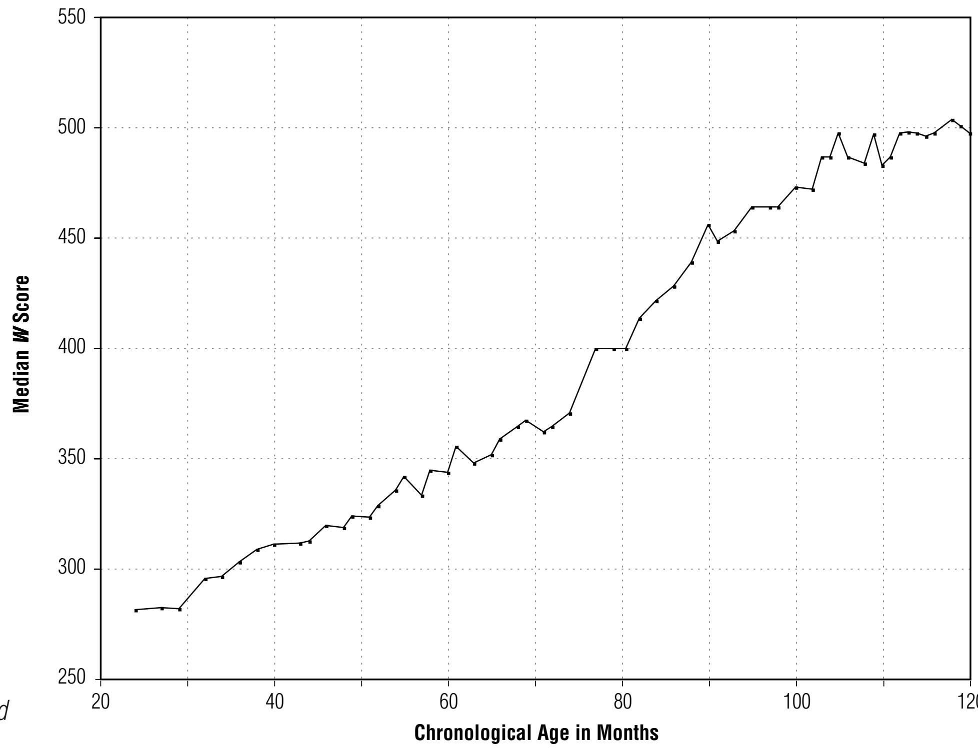



***Figure 3.***
*Plot of select (ages 20 to 120 months only) WJ III Letter-Word Identification age/*W *score sorted block values (from Step 3).*

As Figure 3 shows, although the sample values demonstrate a consistent developmental trend, there is "noise" or "bounce" in the trend due to the aforementioned sampling error. To remove the error from the sample-based data, special polynomial curve-fitting, software-based procedures are employed to produce a "smoothed" solution that best approximates the population REF *W* parameter values (McGrew & Wrightson, 1997; Woodcock, 1994). This process is also repeated for the sample-based standard deviations.[9] Figure 4 presents the result of the polynomial curve-fitting procedures when applied to the Letter-Word Identification data points presented in Figure 3. The smoothed curve provides the normative REF *W* values used in the derivation of WJ III scores (e.g., age/grade equivalents, RPIs, SSs, PRs).[10]

---

[8] For illustrative purposes, age/REF *W* block data points are presented only for subjects ages 20 to 120 months in Figure 3. In practice, the age/REF *W* curves are plotted across the complete age range of the norms for a test.

[9] Additional sources that provide detailed explanations of norm construction via curve-fitting procedures can be found in Daniel (1997), Gorsuch and Zachary (1985), McGrew and Woodcock (2001), McGrew and Wrightson (1997), and Woodcock (1994).

[10] The smoothed norm curve in Figure 4 is illustrative and is not necessarily the final WJ III age norm curve used for the Letter-Word Identification test.

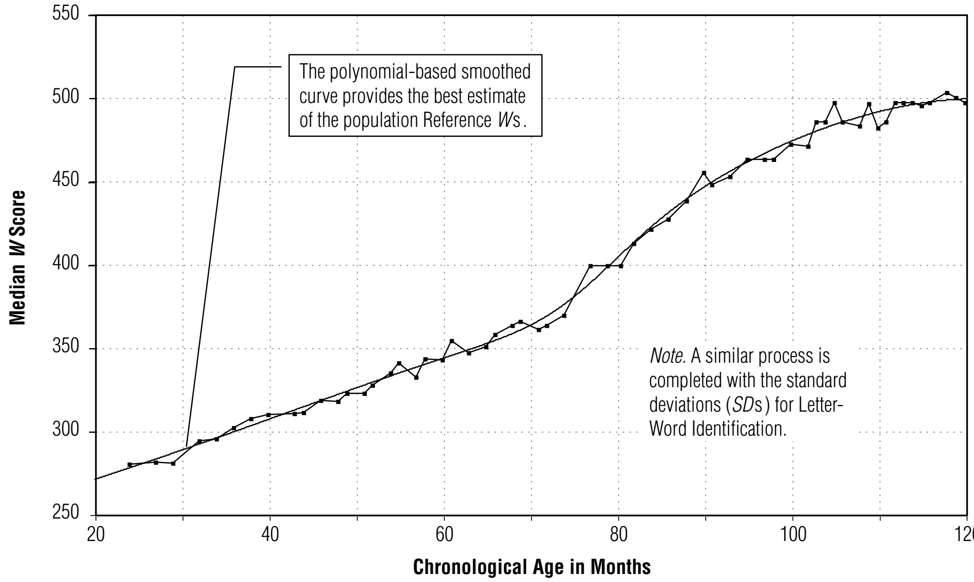



***Figure 4.***
*Smoothed polynomial curve solution for raw age/*W *score Letter-Word Identification sample-based data presented in Figure 3.*

## Construction of WJ III NU Norms: Letter-Word Identification Example

As described above, the use of polynomial curve-fitting procedures in the calculation of the smoothed REF *W* scores (and *SD*s) for all WJ III measures provided, at the time the WJ III was developed and published, norms based on the then-established, state-of-the-art statistical population estimation procedures for calculating derived scores (Daniel, 2007; Gorsuch, & Zachary, 1985; Woodcock, 1994). However, these traditional procedures still did not allow for the recognition of the degree of uncertainty that underlies the raw data points used in the norm curve-fitting procedures. For the calculation of the WJ III NU norms, it was determined that the certainty of the raw data points used to generate norm curves (see Figure 4) could be estimated. This in turn would allow for the incorporation of parameter estimate certainty into the selection of the optimal norm curve solution for all measures via the use of a statistical technique known as the bootstrap procedure.

The bootstrap sampling procedure (Efron & Tibshirani, 1993) is a method for assigning measures of accuracy to statistical estimates. According to the *APA Dictionary of Psychology* (VandenBos, 2007), bootstrap is "a computational method for estimating the precision of an estimate of a (statistic) parameter. A random sample of *n* observations is taken, and from this a number of other samples of equal size are obtained by sampling with replacement" (p. 129). Bootstrap sampling procedures can be used to estimate the uncertainty of a statistic via the provision of a bootstrap standard error (confidence band). This feature is useful in estimating the variability and possible bias in sample statistics, in this case, the sample data used for constructing test norms.

The bootstrap method works by constructing an empirical distribution of a statistic calculated for a sample of subjects drawn from a population. The variability of the statistic within this distribution can be interpreted as a range in which the true value of the statistic would fall if the entire population were to be measured. In simple terms, the bootstrap process, when applied to the calculation of the age/*W* score data points used for curve fitting (see Figure 3), produces a confidence interval/band around the plotted data point, much like the standard error or confidence band clinicians use to bound individual test scores. Repeatedly taking resamples from the obtained norming sample and recalculating the desired statistic constructs the empirical distribution of the statistic or statistics for each resample. As described previously in the WJ III example, one of the desired statistics is the collection of REF *W* values for each successive ordered group of 50 subjects.

In the case of the WJ III NU norm calculation procedures, 250 resamples of the norming sample subjects were taken.[11] A resample is a sample, with replacement, of the same size as the norming sample. Imagine that, for Letter-Word Identification, each subject's age and *W* score are printed on a Ping-Pong™ ball (see Figure 5). All 8,648 balls are placed into a Ping-Pong ball selecting machine (the cylinder in Figure 5). One of the balls (subjects) is randomly selected, the chronological age and *W* score for this selected subject are recorded, and the ball is then thrown back into the machine (replacement) before selecting another one. This process is repeated 8,648 times and ends up producing one resample. Similar to the process described previously for the WJ III Letter-Word Identification norm calculation procedures, Steps 1 and 2 are completed, which results in median chronological ages and *W* scores for each of the age-sorted blocks of 50 subjects. This process (select 8,648 balls, sort records, divide into groups, calculate statistics) is then repeated 250 times. Figure 5 summarizes these steps.

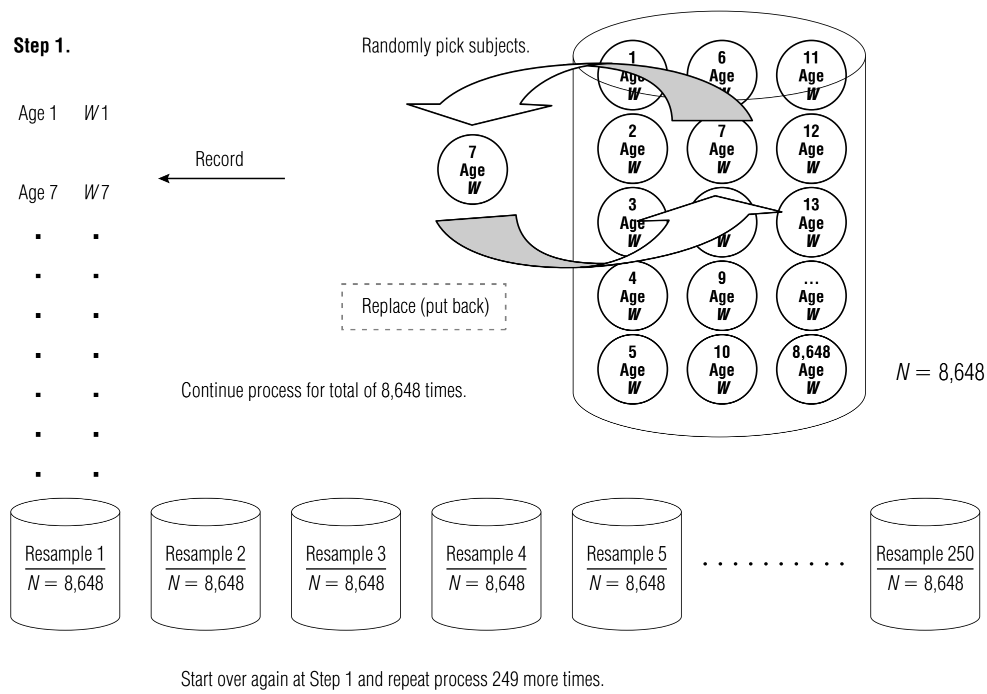



***Figure 5.*** *Creation of 250 WJ III NU Letter-Word Identification resamples via random selection of subjects with replacement (bootstrap method).*

[11]According to Efron and Tibshirani (1993), 50 bootstrap resamples are often sufficient to provide accurate estimates of the standard error of the statistic. With each increase in number of bootstrap resamples, the amount of improvement in the statistical estimates becomes less. A total of 250 resamples were selected to ensure a high degree of confidence in the estimates of the standard error of the block-level statistics.

At this stage, 250 paired values (median chronological age and median *W* score) exist that were calculated for the first (youngest) group in each resample, 250 paired values for the next-youngest group, and so on, up to the 250 paired values for the oldest group of 50 subjects (see Figure 6). This process produces an empirical distribution of 250 values of the statistics for the youngest group of subjects across resamples, for the next youngest group, and so on up to the oldest group. From each of those empirical distributions, the 25th and 75th percentile of each statistic is calculated. The range of sample statistics between the 25th and 75th percentile represents the middle 50% of the generated sample statistics. This window or band provides an empirical estimate of the degree of certainty in the sample statistics that will be used for norm curve generation. If a line is drawn from the point defined by the 25th percentile of the chronological age and the 25th percentile of the median *W* score to the point defined by the 75th percentile of each respective statistic, the result, for each of the age-sorted blocks of subjects, is a “stick” or “window” through which smoothed norm curves are fit. That is, instead of fitting norm curves to single data points (see Figure 3), norm curves are now fit to confidence band windows.

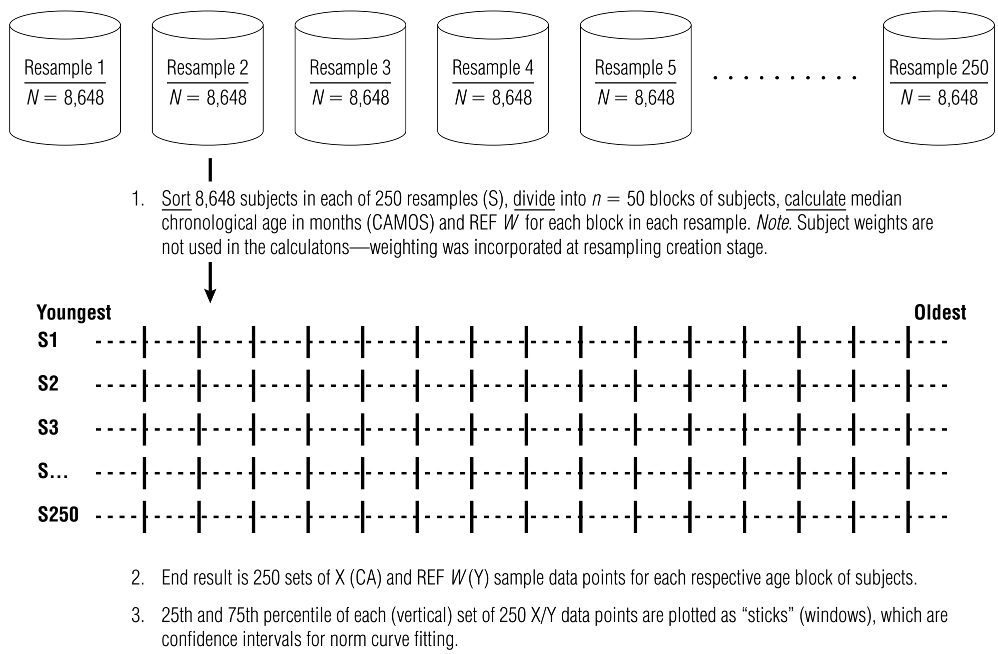



***Figure 6.***
*Calculation of bootstrap-generated sample statistic (see Figure 5) confidence band windows (25th to 75th percentile).*

Aside from the generation of bootstrap resamples, the WJ III and WJ III NU norm development procedures differ in the use of the subject weights. In the case of the WJ III NU resampling procedure, it is not necessary that each subject be as likely to be selected as every other subject. This is where the subject weights come into play. In the WJ III NU, subject weights are converted to selection probabilities, such that subjects with higher weights have a higher chance for selection and inclusion in any given resample. In fact, some subjects are selected many more times than others as this reflects the notion that the subject weights are necessary to balance out the demographic characteristics in the norming sample with respect to the reference population.

There are distinct advantages to using a resampling plan as described here. First, within each resample, the calculation of statistics for each block or group of 50 subjects is simplified because subject weights are no longer part of the calculation. Instead, the subject weights are incorporated in the probability of including a particular subject in each resample. This makes the calculation of more complex statistics (beyond the median *W* score) easily possible.

Second, and more importantly, the norm curve-fitting process involves choosing a path through a series of confidence bands (sticks/windows) instead of a series of single data points. At any given age, there is a range of values (with a known degree of certainty) that might be acceptable smoothed norm REF *W* values. When norm curves are fit to a series of individual data points (WJ III method), there is a tendency to focus on curve solutions that miss as few data points as possible. By fitting a curve through confidence band windows, the uncertainty inherent in the sampling process (as in any sampling process) is acknowledged and visibly observable, which reduces the tendency for norm curves to "chase individual data points," a process that may result in less precise norm curves. This advantage can be seen in Figure 7 where the WJ III NU Letter-Word Identification bootstrap confidence bands are superimposed over the WJ III Letter-Word Identification single data points (from Figure 3).

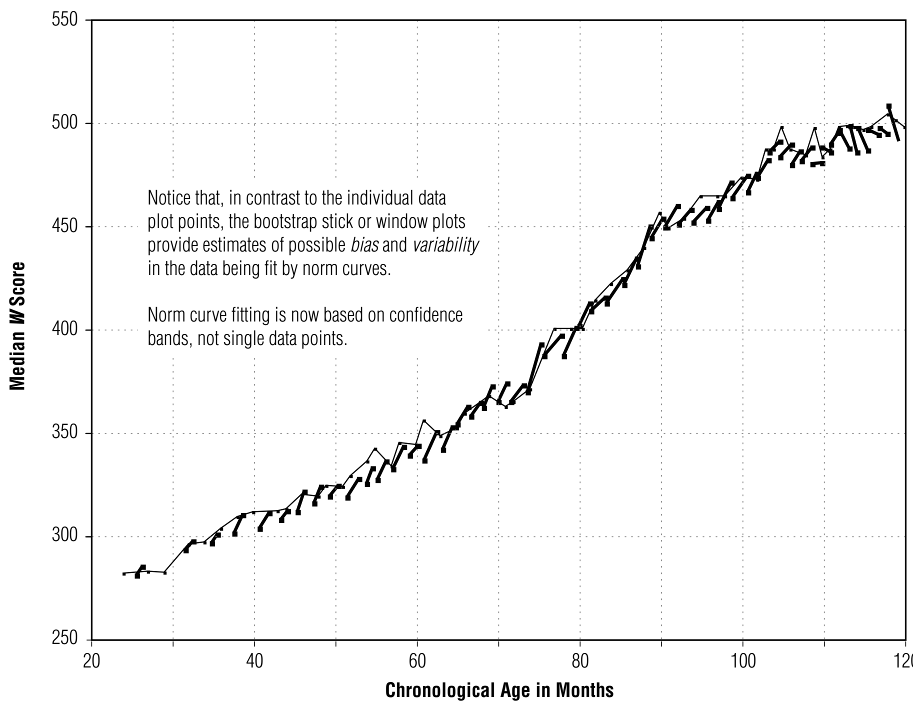


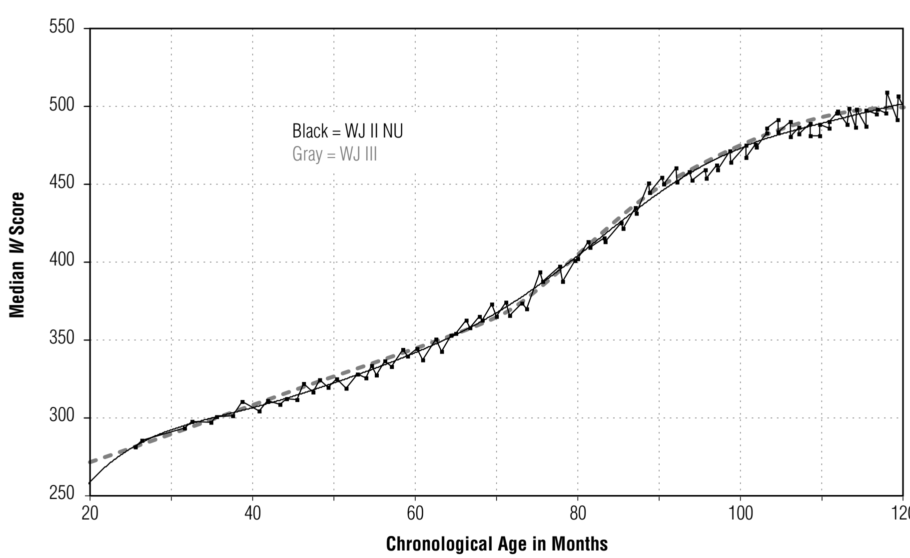

***Figure 7.***
*Comparison of WJ III Letter-Word Identification REF* W *raw data "points" and WJ III NU bootstrap "sticks/windows."*

*Note*. Top and bottom of sticks are slightly offset to allow plotting on the same graph.

The most obvious difference between the single-point WJ III data and the confidence band window-based WJ III NU data in Figure 7 is visible between approximately 35 to 65 months of age. The WJ III NU confidence bands, which are consistently lower than most of the connected single WJ III-derived data, suggest that the WJ III single-point values in this age range are biased upward, and, when a norm curve tracks these upwardly biased data points, it will provide less precise norm curve REF *W* estimates. The evidence is clear from the WJ III NU confidence-band windows that the population REF *W* values in this age range are likely to track lower than the WJ III single-point values. This difference can be seen vividly in Figure 8, where the previously presented WJ III curve solution, although appearing to be an optimal solution in Figure 4 (as also suggested by the various polynomial curve-fitting statistics associated with the solution in Figure 4), does "run high" between 35 to 65 months of age. One can have more confidence, given the quantification of the variability in the range of sample estimates for these specific age blocks (as represented by the sticks/windows), that the more precise norm curve should track lower at these ages. A review of both the illustrative WJ III and WJ III NU curves in Figure 8 demonstrates that the bootstrap method, an improved methodology for norm curve generation (given that it incorporates the uncertainty in the values to be fit in the creation of the norm curves), provides, at a number of age levels, norm curves that either track higher or lower than those based on the older and more traditional WJ III single data-point method. Clearly, greater confidence can be placed in the WJ III NU bootstrap-based norm curves.

Another point where the WJ III and WJ III NU Letter-Word Identification norm curves are noticeably different is from approximately 20 to 30 months. The WJ III point-based curve solution trends are noticeably higher than the WJ III NU curve for this age range. It is well known among applied psychometricians that fitting norm curves at the youngest and oldest ages are the most problematic. When fitting curves to a set of continuous data points surrounded by other data points (e.g., the data points between 60 and 80 months in Figure 3), not only are the specific data-point values within this range used by the curve-fitting algorithms to generate possible solutions, but also information from the data points before 60 months and immediately following 80 months contribute information to the computational algorithm.

However, as can be seen in Figure 3, the first data point has no succeeding or prior data points that contribute information to fitting a proper curve through this first point, the second data point, etc. Conversely, at the oldest ages for a test's norms, the last data point does not similarly benefit from information from data beyond the last data point. As a result, there is considerable uncertainty surrounding the initial starting point (and final ending point) and the shape of the fitted norm curves at the youngest and oldest ages, as norm curves typically are extrapolated slightly beyond the extreme data points available. As a result, curve fitting becomes more art than science at the extremes of norm tables. However, as can be seen in Figure 7, when bootstrap-based confidence bands are the source data for curve fitting, the general trends of the first (and final) sets of sample statistics are more apparent. In the case of the WJ III NU Letter-Word Identification example presented in Figures 7 and 8, greater certainty was placed in a lower WJ III NU norm curve solution between 20 and approximately 36 months. Given this improvement in certainty in fitting the extreme ends of the norm curves, it is not surprising that the largest WJ III/WJ III NU SS differences noted in Tables 1 and 2 are typically at the preschool and older adult levels.




*Note.* Connected data points are actually connected bootstrap sticks/windows (see unconnected sticks/windows in Figure 7).

***Figure 8.***
*Comparison of possible WJ III (gray) and WJ III NU (black) Letter-Word Identification REF* W *norm curves.*

## Illustrative Implication of Letter-Word Identification Example

Given that greater confidence can be placed in the WJ III NU bootstrap-based norm curve solutions (as illustrated in Figures 7 and 8), what do these differences mean for WJ III/WJ III NU score differences? One example follows.

Using the same 35-to 65-month age-span example described previously, a lower (and more precise) set of WJ III NU Letter-Word Identification REF *W* scores will produce the following, given the same obtained WJ III/WJ III NU *W* score (e.g., 325) for an individual.

- Higher age equivalents, as drawing a horizontal line from the Y-axis value corresponding to 325 will intersect the WJ III gray norm curve earlier on the X axis (which, when a vertical line is dropped down to the X axis, the intersection provides the corresponding age equivalent) than its intersection with the lower black WJ III NU norm curve. If the WJ III NU curve is higher than the WJ III norm curve, the converse will hold true (providing lower WJ III NU age equivalents).
- Assuming no difference in the smoothed standard deviations for Letter-Word Identification, for the same Letter-Word Identification *W* score (e.g., 325) for an individual between 35 and 65 months of age, the lower WJ III NU REF *W* norm curve will result in higher measures of relative standing (e.g., SS, PR). For example, as can be seen in Figure 8, the normative or average REF *W* score for an individual with a chronological age of 50 months is approximately 325 (WJ III) and 312 (WJ III NU), respectively.[12] With an obtained *W* score of 325, the subject is "on norm" or average for his or her age when using the WJ III norm curve, which corresponds to a standard score (SS) of 100 and percentile rank (PR) of 50. Conversely, the same obtained *W* score is +13 points higher than the WJ III NU REF *W* value of 312, which indicates the subject is performing above the norm. The exact SS (>100) and PR (>50) for this later example depend on the *SD* associated with scores that are above the REF *W* for 50-month-old individuals (see the WJ III Technical Manual [McGrew & Woodcock, 2001] for an explanation of calculation of measures of relative standing for scores above and below the normative REF *W* score for each age). The converse holds for situations where WJ III NU REF *W* norm curves are higher than the corresponding WJ III norm curve—the subjects, given the same *W* score, would obtain lower SS and PR scores.

It is important to remember that derived metrics of relative standing (e.g., SS, PR) are based on the calculation of the standard *z*-score formula where:

$$z = (\text{observed } W - \text{REF } W) / SD \qquad (1.1)$$

Using this formula, it can be seen that individuals who obtain the same observed *W* score on the Letter-Word Identification test may receive different WJ III/WJ III NU relative metric scores (e.g., SS, PR) as a function of either: (a) higher or lower smoothed WJ III NU REF *W* norm curves, (b) higher or lower smoothed WJ III NU *SD* norm curves, and/or (c) the possible interaction of both a and b.

---

[12]REF *W* values for specific ages are found by ascertaining the *W* score on the Y axis that corresponds to each chronological age value on the X axis of the fitted norm curve (the intersection of the X/Y values on the fitted curve).

In conclusion, significant enhancements in the procedures used to calculate the raw material used for norm curve fitting is one of the major differences between the WJ III and WJ III NU norms. The WJ III norms were not wrong. They were calculated using the then best practice set of procedures typically used by test developers at the time. The WJ III NU has again "raised the bar" for the applied psychometric technology employed in the development of individually administered norm-referenced test instruments.

The use of bootstrap resampling procedures allowed for the incorporation of estimates of uncertainty and potential bias in the sample data used in the calculation of the WJ III NU norms. As a result, the WJ III NU norms represent a technological advancement in the development of test norms. The WJ III NU norms and resultant scores, when compared to the WJ III, are more precise estimates of an individual's tested performance. WJ III/WJ III NU REF *W* and *SD* normative curves may or may not have changed, depending on the specific test or cluster, the developmental status of the tested individual (age or grade), and/or the type of norm bases used (age or grade norms). Thus, it is not possible to provide test/cluster-specific or age/grade-specific rules of thumb regarding potential WJ III/WJ III NU score differences. WJ III/WJ III NU score differences depend on the interaction of the factors above, although users should recognize that "different" does not imply that something is, or was, "wrong." The WJ III NU bootstrap-based norm development procedures provide users with derived scores in which they can place greater confidence.

### What WJ III/WJ III NU scores should be used to track and compare test performance for individuals across time? What is the suggested best practice for tracking student growth from the WJ III to the WJ III NU?

During the transition from the WJ III to the WJ III NU users may encounter situations where examinees will have prior WJ III test results based on the WJ III norms (e.g., 3-year special education reevaluations). If one of the intended purposes of the pretest/posttest comparison is to track growth from the older to the newer examination, examiners should heed a simple principle—comparing the WJ III test scores across time must be based on the same set of norms. Why? Because it is important to remove the effect of possible WJ III/WJ III NU norm differences as a potential confound in measuring growth across time.

If an examinee has prior WJ III-based scores available as the premeasure, the new WJ III administration should be scored both with the WJ III norm-based software *and* the WJ III NU norm-based software. For pretest/posttest growth score comparisons, the first (pre) WJ III norm-based test scores (A) should be compared to the second (post) WJ III norm-based test scores (B), even though a parallel set of WJ III NU scores (C) will also be available. The proper pretest/posttest comparison is to compare the set A and B scores. Interpretations of A to B similarities and differences should focus only on relative change and growth across time. The B set of scores should *not* be used for any other interpretations or decisions (e.g., eligibility, classification, diagnosis). If new decisions are to be made about an examinee's current level of functioning, diagnoses, or classification, then the C set of scores (those based on the WJ III NU norms) should be used. The C scores represent the better and most contemporary norm bases for which to evaluate an individual's performance.

Subsequently, if the same individual is then tested for a third time (D), growth/change comparisons should be made between WJ III NU (C) and WJ III NU (D) scores. The A to B comparison should be viewed as a temporary transition bridge for evaluating growth from the WJ III to the WJ III NU norms.

One additional pretest/posttest comparison possibility might be considered, although the feasibility is less probable and it is *not* the recommended best practice. Instead of comparing A and B scores, if an examiner has access to the original test record information (all raw scores, dates, etc.), it would be technically feasible to rerun the original A performance with the WJ III NU software. In a sense, this is a retroactive scoring of WJ III norm-based performance with the WJ III NU norms. Let's call this set of scores A2. Then, given that the norm basis is the same, it would be possible to evaluate growth or change via an A2 to B score comparison. That is, the current and past WJ III scores are both placed on the common WJ III NU norm basis. Although technically sound for evaluating growth and change, the retroactive rescoring of a previously administered WJ III may be fraught with potential complications. If the A2 scores are markedly different from the original A scores, questions might be asked about prior decisions and recommendations. Such post-hoc "Monday morning quarterbacking" raises a number of potential unknown and complicated issues. It is therefore recommended that if this A2 to B strategy is employed, the A2 scores (old WJ III test performance rescored with WJ III NU norms) *not* be used to retroactively change or question prior decisions, diagnoses, or classifications. The original A scores were the best set of WJ III scores available at the time decisions were made. Previous WJ III norm-based decisions should not be second guessed via the use of computer software scoring.

# Summary

The WJ III Normative Update is a recalculation of the original WJ III normative data based on the final U.S. Bureau of Census statistics, which were released in 2005. In contrast, the original WJ III norms were based on year 2000 census projections that were published in 1996. Changes to the U.S. population demographics provided the initial impetus for an investigation of how such changes would be reflected in a recalculation of the WJ III norms.

The U.S. population had changed sufficiently to make noticeable differences in the resultant subject weights, and, in turn, the scores a subject would obtain on the WJ III NU. In general, many (but not all) of the resulting WJ III NU scores will be lower, particularly at the youngest and oldest age groups. Examiners can use Tables 1 and 2 of this bulletin to obtain information on the typical standard score differences that may be anticipated when updating from the WJ III to the WJ III NU norms.

For development of the WJ III NU norms, greater certainty of the raw data points used to generate norm curves was introduced via the use of a statistical technique known as the bootstrap procedure—a method of assigning measures of accuracy to sampling data. The combined effect of this technique with the updated subject weights provides a more current and accurate comparison of an individual's scores to the U.S. population.

The *Ethical Principles of Psychologists and Code of Conduct of the American Psychological Association* (APA, 2002) suggests that assessments, intervention decisions, and recommendations should be based on an instrument's most current norms (Standard 9.08). This principle suggests that examiners should update to the WJ III NU norms as soon as practicable. However, as described in this bulletin, in limited instances where growth and change across time is being measured, the original WJ III norms should be used as the postchange measure, but only for that purpose.

# References


American Psychological Association (APA). (2002). *Ethical principles of psychologists and code of conduct.* Washington, DC: Author.

Daniel, M. H. (2007). Test Norming [On-line]. Available: http://ags.pearsonassessments.com/assessments/norming.asp

Day, J. C. (1996). *Population projections of the United States by age, sex, race, and Hispanic origin: 1995 to 2050*. U.S. Bureau of the Census, Current Population Reports, (pp. 25–1130). Washington, DC: U.S. Government Printing Office.

Efron, B. & Tibshirani, R. (1993). *An introduction to bootstrap*. New York: Chapman and Hall.

Gorsuch, R. L., & Zachary, R. A. (1985). Continuous norming: Implication for the WAIS-R. *Journal of Clinical Psychology, 41*(1), 86–94.

Landphair, T. (2004, May). Census projections: U.S. population will swell to 420 million by 2050. Retrieved March 2, 2007, from http://voanews.com/english/archive/2004-05a-2004-05-04-34-1.cfm.

McGrew, K., Werder, J., & Woodcock, R. W. (1991). *The WJ-R Technical Manual*. Rolling Meadows, IL: Riverside Publishing.

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). Technical Manual. *Woodcock-Johnson III Normative Update.* Rolling Meadows, IL: Riverside Publishing.

McGrew, K., & Woodcock, R. W. (2001). Technical Manual. *Woodcock-Johnson III.* Rolling Meadows, IL: Riverside Publishing.

McGrew, K. S., & Wrightson, W. (1997). The calculation of new and improved WISC-III subtest reliability, uniqueness, and general factor characteristic information through the use of data smoothing procedures. *Psychology in the Schools, 34*(3), 181–195.

Schrank, F. A., & Woodcock, R. W. (2007). WJ III Normative Update Compuscore and Profiles Program (Version 3.0) [Computer software]. *Woodcock-Johnson III.* Rolling Meadows, IL: Riverside Publishing.

U.S. Census Bureau. (2005). Census 2000 Summary File 3 United States. Washington, DC: Author.

VandenBos, G. (2007). *APA dictionary of psychology.* Washington, DC: American Psychological Association.

Woodcock, R. W. (1994). Norms. In R. J. Sternberg (Ed.), *The encyclopedia of intelligence.* New York: Macmillan.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III.* Rolling Meadows, IL: Riverside Publishing.

Woodcock, R. W., McGrew, K. S., Schrank. F. A., & Mather, N. (2001, 2007). *Woodcock-Johnson III Normative Update.* Rolling Meadows, IL: Riverside Publishing.

Woodcock, R. W., Schrank, F. A., Mather, N., & McGrew, K. S. (2007). *Woodcock-Johnson III Tests of Achievement Form C/Brief Battery.* Rolling Meadows, IL: Riverside Publishing.

Riverside Publishing
A HOUGHTON MIFFLIN COMPANY

3800 Golf Road, Suite 100
Rolling Meadows, IL 60008-4015

**800.323.9540**
**www.woodcock-johnson.com**

**1051547**