Variability in Wechsler Adult Intelligence Scale-IV Subtest Performance Across Age

Nick M. Wisdom\textsuperscript{1,2}, Joseph Mignogna\textsuperscript{1,3}, Robert L. Collins\textsuperscript{4,5,*}

\textsuperscript{1}Mental Health Care Line, Michael E. DeBakey VA Medical Center, Houston, TX, USA
\textsuperscript{2}Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, USA
\textsuperscript{3}VA HSR&D Houston Center of Excellence, Michael E. DeBakey VA Medical Center, Houston, TX, USA
\textsuperscript{4}Neurology Care Line, Michael E. DeBakey VA Medical Center, Houston, TX, USA
\textsuperscript{5}Department of Neurology, Baylor College of Medicine, Houston, TX, USA

*Corresponding author at: Neurology Care Line, Michael E. DeBakey Veterans Affairs Medical Center, Houston, TX, USA. Tel.: +713-794-8939; fax: +713-794-7786.
E-mail address: robert.collins3@va.gov (R.L. Collins).

Accepted 15 March 2012

Abstract

Normal Wechsler Adult Intelligence Scale (WAIS)-IV performance relative to average normative scores alone can be an oversimplification as this fails to recognize disparate subtest heterogeneity that occurs with increasing age. The purpose of the present study is to characterize the patterns of raw score change and associated variability on WAIS-IV subtests across age groupings. Raw WAIS-IV subtest means and standard deviations for each age group were tabulated from the WAIS-IV normative manual along with the coefficient of variation (CV), a measure of score dispersion calculated by dividing the standard deviation by the mean and multiplying by 100. The CV further informs the magnitude of variability represented by each standard deviation. Raw mean scores predictably decreased across age groups. Increased variability was noted in Perceptual Reasoning and Processing Speed Index subtests, as Block Design, Matrix Reasoning, Picture Completion, Symbol Search, and Coding had CV percentage increases ranging from 56\% to 98\%. In contrast, Working Memory and Verbal Comprehension subtests were more homogeneous with Digit Span, Comprehension, Information, and Similarities percentage of the mean increases ranging from 32\% to 43\%. Little change in the CV was noted on Cancellation, Arithmetic, Letter/Number Sequencing, Figure Weights, Visual Puzzles, and Vocabulary subtests (<14\%). A thorough understanding of age-related subtest variability will help to identify test limitations as well as further our understanding of cognitive domains which remain relatively steady versus those which steadily decline.

Keywords: Statistical methods; Norms/normative studies; Intelligence; Everyday functioning; Developmental and learning disabilities; Assessment

Introduction

An effective neuropsychological evaluation is contingent on recognizing several important mitigating factors that influence the interpretation of collected test scores. In the evaluation of older adults, perhaps the most important issue is the differentiation between normal, age-related changes and pathologic brain dysfunction (Miller, Myers, Prinzi, & Mittenberg, 2009; Nadler, Mittenberg, DePiano, & Schneider, 1994). Normal, non-pathological aging is associated with a number of physiologic changes and neurological declinations. The most common findings include the enlargement of the ventricles, cortical sulci, and subarachnoid space; gross cortical, hippocampal, and temporal lobe atrophy; hyperintensity in the white matter and basal ganglia; and arteriosclerosis (Drayer, 1988; Scahill et al., 2003). However, the widespread belief that neuronal loss is associated with normal aging has recently been criticized and there is some support that decreases in brain matter are more the result of reduced gray matter volume (Abe et al., 2008; for a review, see Keller, 2006).

These age-associated reductions in gray matter have been shown to negatively impact cognitive functioning in healthy, older adults (Zimmerman et al., 2006). Regardless of etiology, cognitive change as a normal process of advancing age is well documented in the literature. More specifically, declines in executive functioning (Royall, Palmer, Chiodo, & Polk, 2004; West &
Much of the aforementioned research was conducted using the most widely used and researched measure of intelligence—the various iterations of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955, 1981, 1997, 2008a). At its core, the Wechsler scales have always been comprised of point scales, in that an overall Intelligence Quotient (IQ) is based on the summation of “points earned” across various subtests. As subtests are grouped together to form both factor scores and an IQ, an underlying assumption exists that each subtest is accorded the same weight as the others (Matarazzo, 1972). As a result, tests with lower reliabilities (e.g., object assembly) were included in the overall IQ because intelligence was originally considered “assortative” rather than hierarchical in that each subtest adds unique information to the overall IQ (Matarazzo, 1972). Over time, much more emphasis has been placed on individual subtests and their semantic combinations (i.e., index scores). It is now commonly held that there is a real benefit in being able to compare an individual’s performance across subtests and factors and that many of these scores can be considered indicators of neurological integrity (Ruff, Allen, Farrow, Niemann, & Wylie, 2002) are prominent with aging. Given these findings, it is not surprising that age significantly influences many of the tests utilized to measure intelligence (Kaufman, 1990). 

Ardila (2007), working backward through the normative tables, examined age-related changes in variability on WAIS-III subtests by calculating the “percentage of the mean” for each subtest. The percentage of the mean, more formally known as the coefficient of variation (CV), is simply a ratio of the standard deviation/mean × 100 (Bartlett, 1946; Hendricks & Robey, 1936; Pearson, 1897; Yablokov, 1974). The CVs were previously reported for the Wechsler–Bellevue index scales (Verbal, Performance, Full Scale; Matarazzo, 1972), but Ardila (2007) was the first to examine variability and score dispersion at the subtest level for the WAIS-III. Unfortunately, the CV is typically overlooked in the behavioral sciences despite having two very attractive properties. First, by expressing the variability as a percentage of the mean, more information about actual variability is conveyed; that is to say, a standard deviation score alone is not meaningful without reference to the mean. In turn, this enables researchers and clinicians to subsequently compare the CVs within and between populations (Lande, 1977). Using the CV methodology, Ardila demonstrated that mean scores on the WAIS-III performance generally decline across age cohorts, while the heterogeneity for most subtests increases greatly. However, the exact percentage increase/decrease in the CV for each subtest is difficult to discern from Ardila’s study as he chose to express this change as a “ratio percentage” instead of the percentage change.

As a result of several factor-analytic models, the WAIS-IV has implemented a number of changes and now has 15 subtests across four index scales, eschewing the traditional dual IQ (Verbal and Performance) index score structure from previous iterations of the WAIS (for a review, see Wechsler, 2008b). The WAIS-IV has retained all the Verbal Comprehension Index (VCI) subtests from the WAIS-III, although Comprehension is now supplemental. Likewise, the Working Memory Index (WMI) subtests remain unchanged, although a new sequencing task has been added to the traditional Forward/Backward Digit Span. The WAIS-III Perceptual Organization Index has been renamed for the WAIS-IV as the Perceptual Reasoning Index (PRI); Block Design and Matrix Reasoning have been retained and Picture Completion is now optional. In addition, two new subtests, Visual Puzzles and Figure Weights (supplemental), have been added to the PRI as measures of non-verbal abstract reasoning and quantitative/analogue reasoning, respectively. Finally, the Processing Speed Index (PSI) has added the supplemental Cancellation task alongside core Coding and Symbol Search subtests. Picture Arrangement and Object Assembly from the WAIS-III previously only loaded on the omnibus Full-Scale IQ and both have been discontinued for the WAIS-IV. For a full review of all changes, including subtest modifications, sample characteristics, and reliability/validation statistics, see the WAIS-IV technical manual (Wechsler, 2008b).

Given the number of substantial changes to the test content and structure found in the WAIS-IV, further evaluation of age-related changes in subtest variability is warranted. The purpose of this study was to calculate the CVs in all of the WAIS-IV subtests across each age grouping. It was hypothesized that mean scores would decrease with age, particularly within the PSI and PRI, while CV scores would amplify, supporting the notion of increased intellectual heterogeneity with aging. Also, the addition of the sequencing task to the Digit Span subtest, along with normative data tabulated for forward/backward, allowed for a unique examination of Digit Span performance across age cohorts. Finally, the results of the current study were compared with Ardila’s (2007) research to better capture changes between WAIS-III and WAIS-IV.
Methods

Normative data found in the WAIS-IV Administration and Scoring Manual (Tables A.1 and C.1; Wechsler, 2008a) were used to determine mean raw scores and CVs (i.e., dispersion scores) within age groups for each WAIS-IV subtest. The procedure described by Ardila (2007) was utilized to determine these scores. Specifically, within each age group, the raw score corresponding to each subtest scaled score = 10 provided the mean raw score for each subtest. When a range was provided for the corresponding raw score (i.e., 71–75), the mean of the range was used. Then, to determine the raw standard deviation within age groups for each subtest, the raw scores corresponding to a scaled score = 7 (1 SD below the mean) and a scaled score = 13 (1 SD above the mean) were separately subtracted from the mean, added together, and then divided by 2. The CV for each subtest by age was determined by dividing the standard deviation by the mean and multiplying by 100. This provides the percentage of the mean represented by the standard deviation (Hendricks & Robey, 1936), and thus information about the magnitude of variability represented by each standard deviation, as well as a standardized score of variability that allows for inter- and intra-age group comparisons.

The WAIS-IV was standardized and normalized using 2,200 American examinees divided into 13 age groups: 16–17, 18–19, 20–24, 25–29, 30–34, 35–44, 45–54, 55–64, 65–69, 70–74, 75–79, 80–84, and 85–90. The ratio of different ethnicities and gender were proportionally consistent with census data within each corresponding age group. The sample was drawn from four different geographical regions in the USA (West, Midwest, South, and Northeast) and stratified according to five education levels based on the number of years of school completed: ≤8, 9–11, 12, 13–15, and ≥16. A detailed description of the sample can be found in the WAIS-IV Technical and Interpretive Manual (Wechsler, 2008b).

Results and Discussion

The raw mean, standard deviation, and CV scores in each age range for all subtests on the WAIS-IV can be found in Tables 1 (VCI), 2 (PRI), 3 (PSI), and 4 (WMI). In addition, Table 5 presents the highest and lowest mean scores, along with the associated mean and CV score percentage change, within each subtest. In other words, the data presented in Table 5 can be used to determine which subtests fluctuate the most with increasing age and which remain stable. More specifically, subtests where the CV remains constant across the lifespan while the mean decreases suggest that the underlying cognitive construct really is declining for most individuals (Ardila, 2007). For example, Visual Puzzles mean scores steadily decline with age, but the CV (i.e., intragroup variability) remains stable in the older age groups. Clinically, this indicates that most people will demonstrate a decline in this measure with age. Compare this with subtests where the CV increases greatly in the context of the mean score decline (e.g., Digit Span Sequencing). Because the older age groups are more heterogeneous, statements regarding an individual’s strength or weakness can be made with less confidence because the spectrum of “normal” is wider. Finally, Table 6 presents the percentage change in CV scores for each subtest for both WAIS-III and WAIS-IV. It is important to note that Ardila (2007), when examining the change in the mean and the associated CV for each subtest across the lifespan, originally calculated a ratio percentage [(lowest mean/highest mean) × 100]. Using the data presented in Ardila’s manuscript, we have

Table 1. Calculated raw scores of the WAIS-IV VCI subtests

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Similarities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>23.0</td>
<td>23.5</td>
<td>24.5</td>
<td>25.0</td>
<td>25.5</td>
<td>26.0</td>
<td>26.0</td>
<td>25.5</td>
<td>24.5</td>
<td>23.5</td>
<td>22.5</td>
<td>21.5</td>
<td>20.5</td>
</tr>
<tr>
<td>SD</td>
<td>5.3</td>
<td>5.3</td>
<td>5.5</td>
<td>5.3</td>
<td>5.5</td>
<td>5.8</td>
<td>5.8</td>
<td>6.0</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
</tr>
<tr>
<td>CV</td>
<td>22.8</td>
<td>22.3</td>
<td>22.4</td>
<td>21.0</td>
<td>21.6</td>
<td>22.1</td>
<td>22.1</td>
<td>23.5</td>
<td>25.5</td>
<td>26.6</td>
<td>27.8</td>
<td>29.1</td>
<td>30.5</td>
</tr>
<tr>
<td><strong>Vocabulary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>30.0</td>
<td>31.0</td>
<td>33.0</td>
<td>34.5</td>
<td>36.0</td>
<td>37.5</td>
<td>38.5</td>
<td>38.5</td>
<td>37.5</td>
<td>36.5</td>
<td>35.5</td>
<td>34.5</td>
<td>33.0</td>
</tr>
<tr>
<td>SD</td>
<td>10.3</td>
<td>10.3</td>
<td>10.3</td>
<td>10.8</td>
<td>11.3</td>
<td>11.5</td>
<td>11.8</td>
<td>11.8</td>
<td>11.5</td>
<td>12.0</td>
<td>12.0</td>
<td>12.3</td>
<td>12.5</td>
</tr>
<tr>
<td>CV</td>
<td>34.2</td>
<td>33.1</td>
<td>31.1</td>
<td>31.2</td>
<td>31.3</td>
<td>30.7</td>
<td>30.5</td>
<td>30.5</td>
<td>30.5</td>
<td>30.7</td>
<td>32.9</td>
<td>33.8</td>
<td>35.5</td>
</tr>
<tr>
<td><strong>Information</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>13.0</td>
<td>13.0</td>
<td>13.0</td>
<td>13.0</td>
<td>13.0</td>
<td>13.5</td>
<td>14.5</td>
<td>14.5</td>
<td>14.5</td>
<td>14.5</td>
<td>13.5</td>
<td>12.5</td>
<td>11.5</td>
</tr>
<tr>
<td>SD</td>
<td>4.8</td>
<td>4.8</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.3</td>
<td>5.3</td>
<td>5.3</td>
<td>5.8</td>
<td>5.8</td>
<td>5.8</td>
<td>5.8</td>
<td>5.8</td>
</tr>
<tr>
<td>CV</td>
<td>36.5</td>
<td>36.5</td>
<td>38.5</td>
<td>38.5</td>
<td>38.5</td>
<td>38.9</td>
<td>39.7</td>
<td>39.7</td>
<td>39.7</td>
<td>42.9</td>
<td>42.9</td>
<td>46.0</td>
<td>50.0</td>
</tr>
<tr>
<td><strong>Comprehension (optional)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>22.5</td>
<td>22.5</td>
<td>23.5</td>
<td>23.5</td>
<td>24.5</td>
<td>25.5</td>
<td>25.5</td>
<td>25.5</td>
<td>24.5</td>
<td>23.5</td>
<td>22.5</td>
<td>21.5</td>
<td>20.5</td>
</tr>
<tr>
<td>SD</td>
<td>5.8</td>
<td>6.0</td>
<td>6.3</td>
<td>5.8</td>
<td>6.3</td>
<td>5.8</td>
<td>6.3</td>
<td>6.3</td>
<td>6.5</td>
<td>6.3</td>
<td>6.3</td>
<td>6.0</td>
<td>6.3</td>
</tr>
<tr>
<td>CV</td>
<td>25.6</td>
<td>26.7</td>
<td>26.6</td>
<td>24.5</td>
<td>25.5</td>
<td>22.5</td>
<td>24.5</td>
<td>25.5</td>
<td>27.7</td>
<td>27.8</td>
<td>29.1</td>
<td>29.3</td>
<td>32.1</td>
</tr>
</tbody>
</table>

*Note: CV = coefficient of variation.*
instead calculated the more intuitive percentage increase/decrease for both the mean and CV for each subtest \[(\text{amount of change/highest mean or associated CV}) \times 100\] and presented them in Table 6.

**Verbal Comprehension Index**

In general, all raw mean scores followed a similar and predictable pattern, increasing from younger to middle age before steadily declining into older adulthood. All the VCI subtests decline by about 20% when measuring the percentage change from the highest to the lowest mean scores; however, with the exception of Vocabulary, the intragroup variability as represented by the CV increases moderately. More specifically, the associated CV increases by a moderate 40% when moving from the highest to the lowest mean scores on the Similarities, Information, and Comprehension subtest. Vocabulary is unique in that it is the only WAIS-IV subtest where the lowest mean scores are not observed in the oldest age group, but rather the youngest (16–17 years). Vocabulary is also distinctive among VCI subtests because there is little intragroup variability in performance, irrespective of age, as the CV only varies 12% when comparing the associated highest and lowest mean scores.
Perceptual Reasoning Index

A general pattern of the mean score emerges when examining the PRI subtests in that all of the mean scores strongly decline with age—the highest and lowest mean scores are found in the youngest and oldest age groups, respectively. Excluding Figure Weights, all the PRI subtest mean scores decline by about 53%. Figure Weights mean scores only decline by 27%,
but it is important to note that normative data for individuals over the age of 70 were not reported in the WAIS-IV manual. Figure Weights would likely follow a similar pattern through older adulthood as all the related PRI subtest mean score declines are comparable in the 65–69 age grouping. Despite the uniform decline in mean scores, a bimodal distribution of CV scores suggests more variability both between and within measures that might otherwise be overlooked by focusing exclusively on changes in mean scores. When measuring the percentage change in CV scores associated with the highest and lowest means, a strong increase is observed on the Block Design (56.1%), Picture Completion (86.6%), and Matrix Reasoning (97.9%) subtests, suggesting increasingly heterogeneous performance within each subsequent age grouping that is masked by the mean. The intragroup variability for Visual Puzzles, however, actually decreases with age by 1.6%. Similarly, Figure Weights increases only slightly by 11.3%. This change cannot be explained by the lack of older adult normative information since the other subtest CV scores greatly surpass this minimal increase before reaching the 65–69 age range.

Processing Speed Index

A strong decline in performance with aging is evident, as the highest and lowest PSI subtest mean scores are found in the youngest and oldest age groups. Symbol Search and Coding exhibit a similar trajectory: mean scores strongly decline with increases in age by about 54%, while associated CVs strongly increase by 81% and 93%, respectively. Cancellation, however, appears to be more resilient to the effects of aging as the mean only declines by 14% with a small 8% increase in intragroup variability. Although the complete performance curve for the Cancellation subtest cannot be calculated as it was not normed in individuals over the age of 70, it is worth noting that the magnitude of mean and CV change for Coding and Symbol Search is moderately larger than Cancellation in the 65–69 age range. The increased intragroup variability noted on these measures is particularly noteworthy as processing speed is a cognitive construct that has been cited as the primary underlying process for loss of cognitive functioning in normal aging (Salthouse, 1996).

Working Memory Index

The pattern of decline from the highest means scores to the lowest is less discernible among the WMI subtests. Letter-Number Sequencing exhibits the least amount of mean score change, declining by 7%, but this appears to be an artifact given the lack of normative data for the oldest age groups; the other WMI subtests have a similar pattern of decline before reaching the 65–69 age range. The remaining WMI subtests demonstrate small to middling mean score decreases with age, ranging from 18.2% (Digit Span Forward) to 33.3% (Digit Span Sequencing). With the exception of Digit Span Sequencing, the associated CV changes are relatively uniform. Intragroup variability actually decreases with advancing age on the Digit Span Forward (−2.2%) and Digit Span Backward (−3.2%) subtests. In addition, minimal CV score increases are observed on the Letter-Number Sequencing (7.7%) and Arithmetic (13.9%) subtests.

In contrast, Digit Span Sequencing appears to be inherently unique in that it demonstrates a very different pattern of score dispersion. Digit Span Sequencing is suggested to be similar to Digit Span Forward and Backward in that both aspects of working memory and mental manipulation are cognitively taxed (Wechsler, 2008b). However, Digit Span Sequencing is thought to theoretically differ from Digit Span Forward (attention efficacy) and Digit Span Backward (transformation of

<table>
<thead>
<tr>
<th>Subtest</th>
<th>WAIS-III (Ardila, 2007)</th>
<th>WAIS-IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High-score age</td>
<td>Low-score age</td>
</tr>
<tr>
<td>Similarities</td>
<td>30–54</td>
<td>85–89</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>45–54</td>
<td>85–89</td>
</tr>
<tr>
<td>Information</td>
<td>45–54</td>
<td>85–89</td>
</tr>
<tr>
<td>Comprehension</td>
<td>45–54</td>
<td>85–89</td>
</tr>
<tr>
<td>Block Design</td>
<td>16–29</td>
<td>85–89</td>
</tr>
<tr>
<td>Matrix Reasoning</td>
<td>16–17</td>
<td>85–89</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>16–44</td>
<td>85–89</td>
</tr>
<tr>
<td>Digit Span, Overall</td>
<td>16–29</td>
<td>85–89</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>45–54</td>
<td>85–89</td>
</tr>
<tr>
<td>Letter-Number Seq.</td>
<td>16–44</td>
<td>85–89</td>
</tr>
<tr>
<td>Symbol Search</td>
<td>18–19</td>
<td>85–89</td>
</tr>
<tr>
<td>Coding/Digit Symbol</td>
<td>18–19</td>
<td>85–89</td>
</tr>
</tbody>
</table>

Notes: WAIS = Wechsler Adult Intelligence Scale; CV = coefficient of variation.
information), in that sequencing is a more demanding task requiring simultaneous activation of phonological and semantic information as a precursor to task comprehension (MacDonald, Almor, Henderson, Kempler, & Andersen, 2001; Werheid et al., 2002). In our analysis of WAIS-IV Digit Span subtests, we found that Digit Span Sequencing appeared to change differentially across age groups, with both Digit Span Forward and Backward raw and CV scores remaining relatively stable, whereas Digit Span Sequencing raw scores decreased by 33% from the youngest to the oldest groups and the CV increased by 87.8%. Creating composite WAIS Digit Span subtest scores (e.g., combining Forward and Backward) has been criticized previously, as both are felt to provide distinct information (Banken, 1985; Reynolds, 1997). This would also appear to be the case with Digit Span Sequencing and the distinction between the Digit Span subtests appears to be amplified across age cohorts.

Comparison of WAIS-III and WAIS-IV

In general, the majority of the WAIS-IV subtests demonstrate similar CV changes as those observed on the WAIS-III, despite a number of material and administration differences. This suggests that either these subtests (e.g., Symbol Search) are psychometrically similar between the WAIS-III and WAIS-IV iterations and/or that the underlying latent construct (e.g., processing speed) is robust to measurement. A few notable exceptions to this pattern include Digit Span, Letter-Number Sequencing, Block Design, and Picture Completion. The variability in overall Digit Span scores among the older age groups is much higher on the WAIS-IV, although this is entirely a result of the additional Sequencing task which was previously not included in the WAIS-III. Likewise, Letter-Number Sequencing appears to vary much less on the WAIS-IV, although as previously noted, this is most likely because it was not normed among individuals over the age of 70 as was previously done on the WAIS-III. Block Design and Picture Completion, however, demonstrate notable volatility. The CV change between the WAIS-III and WAIS-IV indicates that performance among the older age groups on Block Design and Picture Completion greatly increases and decreases, respectively. The reason for this variability is not entirely clear although it does suggest that these subtests have changed considerably and/or the underlying construct of visuospatial perception is being inconsistently measured. Although caution is warranted when clinically trying to transpose any research conducted using previous WAIS iterations onto the WAIS-IV (Loring & Bauer, 2010), particular attention should be directed at studies that have evaluated Block Design and Picture Completion (Ronnlund & Nilsson, 2006; Schoenberg, Scott, Duff, & Adams, 2002).

Conclusion

Healthy aging occurs in conjunction with a decline in many cognitive faculties as measured by the WAIS-IV. However, examining the WAIS-IV subtest mean scores in isolation neglects the variability in performance within each age grouping. A seemingly forgotten, yet valuable way of examining the dispersion of scores is to calculate the CV for each subtest (Hendricks & Robey, 1936; Yablokov, 1974). The CV represents the percentage of the mean score that is accounted for by the spread of 1 SD. Assuming that the standard deviation remains constant, declines in mean scores with aging like those observed on nearly all WAIS subtests necessarily mean increased within-group variability in performance on the subtest. The current study examined the dispersion of scores on the various WAIS-IV subtests across all age ranges included in the normative manual. The findings from this study suggest that there is more heterogeneity in cognitive functioning among older adults than is observed in younger adults. This pronouncement is not particularly surprising as older adults have had significantly more opportunities to experientially shape their cognitive abilities. However, examining only the mean scores creates the false belief that cognition declines uniformly with age when in actuality the intragroup variability amplifies with age.

There are several limitations of this study that should be targeted for improvement in follow-up experiments. Most importantly, this study compares different groups of subjects across different age ranges. This cross-sectional approach often results in a “cohort effect” as each generational group has their own shared temporal life experiences. As a result, the cohort effect may explain some of the differences between age groups rather than real changes in cognitive functioning across the lifespan. Ideally, longitudinal studies should be conducted measuring these cognitive abilities in the same subjects across the lifespan to help improve the internal validity of these findings. Another limitation is that the means and standard deviations represented here are a very close estimation of the observed scores based on the information found in the WAIS-IV manual, but the actual data could not be obtained. Finally, it is important to note that this study only describes and sheds light on the pattern of cognitive decline as measured by the WAIS-IV, but offers no explanation for the changes in intragroup variability observed on many of the subtests. Continued research to help establish the internal validity of these measures in older adults may help explain some of these changes. Studies implementing a double-dissociation design can help differentiate the specificity of these measures for underlying cognitive abilities (Weiskrantz, 1991).
None declared.

References


