Comparison of Predicted-difference, Simple-difference, and Premorbid-estimation methodologies for evaluating IQ and memory score discrepancies

Reid L. Skeel a,*, David Sitzer a, Tammie Fogal a, Janet Wells a, Brick Johnstone b

a136 Sloan Hall, Central Michigan University, Mt. Pleasant, MI 48859, USA
bUniversity of Missouri—Columbia, St. Louis, MO, USA

Accepted 1 April 2003

Abstract

Discrepancies between WAIS-III and WMS-III scores for a group of 39 males and 48 females with a history of TBI were examined using three methodologies: Predicted-difference, Simple-difference, and Premorbid-estimation methods. Overall, the Predicted-difference method tended to classify the fewest individuals as impaired based on statistical rarity of discrepancies (11–16% classified as impaired), while the regression-based Premorbid method tended to classify the fewest individuals as impaired based on clinical rarity of discrepancies (4–8% classified as impaired). Degree of agreement is reported and was substantial. The only comparison between methods to reach statistical significance was the Predicted-difference method classifying subjects as impaired at a higher rate than other methods for Auditory Delayed memory index (Cochran’s $Q = 7.00, P < .05$). Findings suggest a combination of estimates of premorbid functioning and regression-based predicted scores is optimal for interpreting IQ/memory score discrepancies. Clinical implications are discussed.

© 2003 National Academy of Neuropsychology. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Neuropsychological assessment; Wechsler Memory Scale; Premorbid functioning

* Corresponding author. Tel.: +1-989-774-6485; fax: +1-989-774-2553.
E-mail address: reid.skeel@cmich.edu (R.L. Skeel).

0887-6177/02/$ – see front matter © 2003 National Academy of Neuropsychology.
doi:10.1016/S0887-6177(03)00072-6
1. Introduction

The identification of cognitive impairment has long been one of the central tasks for neuropsychology (Schinka & Vanderploeg, 2000). Inherent in this process is deciding whether or not a pattern of scores represents impairment. However, there is no universal agreement among neuropsychologists as to what represents impairment, and there is no standard procedure for calculating impairment (Cimino, 2000). Typically, neuropsychologists use some combination of comparisons between a person’s general level of functioning on current testing and an individual’s estimated level of premorbid functioning. What level of difference qualifies as decline, and what criteria are used to evaluate decline can vary as a function of the measures used and the choices clinicians make to calculate decline.

The WAIS-III–WMS-III Technical Manual (The Psychological Corporation, 1997) outlines two approaches for examining discrepancies between memory and intellectual functioning. These co-normed intelligence and memory scales have provided the opportunity to make direct comparisons between memory and IQ scores allowing for increased accuracy in interpretation of statistically significant differences compared to measures that are not from the same normative sample (Iverson, 2001). When comparisons are made between measures that are normed on different samples, error is introduced into the process due to the potential systematic differences in the normative sample (Cimino, 2000). The WAIS-III–WMS-III Technical Manual (The Psychological Corporation, 1997) notes that comparisons are possible for both index scores and subtest scores. These comparisons may be used to make clinical inferences of memory deficits. In the first approach outlined in the manual (Simple-difference method), difference scores between obtained memory and IQ scores are used to evaluate impairment, while in the second approach (Predicted-difference method) predicted memory scores derived from the standardization sample are calculated according to an individual’s obtained IQ scores. The Predicted-difference method has the advantage of accounting for effects of regression to the mean and measurement error (Bennett & Clarizio, 1988; Braden & Weiss, 1988; Evans, 1992a, 1992b).

As noted in the manual, both approaches may be affected by changes in IQ secondary to neurological changes that may depress both IQ and memory scores. This has the potential to lead to systematic overestimation or underestimation of deficits. Thus, it may be necessary to consider premorbid functioning, particularly when a neurologic event may have impacted IQ scores. It is not clear to what degree interpretation of impairment in a clinical population may be affected by these various interpretive strategies. Techniques developed to analyze discrepancies between IQ and academic achievement (Reynolds, 1990) may be adapted to evaluate differences between premorbid intellectual functioning and memory. In a fashion similar to WAIS–WMS discrepancy analysis, IQ and achievement scores have been compared using simple differences and regression-based predicted scores (Evans, 1992a; Glaub & Kamphaus, 1994; Reynolds, 1990).

While there are theoretical considerations for choosing particular approaches to test score analysis (Lezak, 1995), there has been little empirical comparison of the approaches in a clinical sample. Thus, it is unclear to what degree the various methods may differentially classify particular patient samples as impaired. In a sample of individuals who sustain equivalent declines in memory and IQ, methods based solely upon current obtained scores would tend
to underdiagnose impairment since predicted memory scores based on an already declined IQ would be artificially low. Methods involving premorbid estimation would be less sensitive to these co-occurring declines. The current study compared three methods of interpreting score discrepancies, Simple-difference, Predicted-difference, and a method involving premorbid estimation (Premorbid method) for a sample of individuals who sustained traumatic brain injuries (TBI). Due to co-occurring mild declines in VIQ and memory associated with TBI (Johnstone, Hexum, & Ashkanazi, 1995; Skeel, Johnstone, Schopp, Shaw, & Petroski, 2000), it was hypothesized that regression-based methods involving premorbid estimation would be the most sensitive to decline resulting in the highest rates of individuals being classified as impaired. It was also hypothesized that the Simple-difference method would be least sensitive due to mild VIQ declines in the sample and difficulties inherent in difference scores.

2. Method

2.1. Participants

The sample included 39 males and 48 females referred for neuropsychological evaluation at a Midwestern university medical center who had a history of TBI (see Table 1 for injury severity information). The mean age for the sample was 35.6 (S.D. = 12.2), the mean educational level for the sample was 12.9 years (S.D. = 2.67), and the mean time since injury at the time of testing was 106 months (Mdn = 36). In order to limit the impact of acute effects of TBI, all participants were seen on an outpatient basis with a minimum time since injury of 2 months. Consistent with the demographic make-up of the surrounding community, the sample consisted of 85 Caucasian participants, and 2 African American participants. Additional information concerning SES beyond educational level was not available. Only first neuropsychological evaluations were included in the study. Patients were selected from a larger sample based on four criteria: (1) no history of learning disability or other language disorder; (2) individuals with pathological scores on measures of effort were excluded; (3) individuals completed necessary WAS-III and WMS-III subtests; and (4) information was available concerning date of injury and severity of injury (as measured by loss of consciousness).

Table 1
Mean VIQ and memory scores by injury severity and total sample

<table>
<thead>
<tr>
<th>Injury severity</th>
<th>VIQ</th>
<th>WRAT-III</th>
<th>WMS-AI</th>
<th>WMS-AD</th>
<th>WMS-GM</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC &lt; 1 (N = 42)</td>
<td>97.7 (11.2)</td>
<td>97.2 (10.9)</td>
<td>100.4 (14.1)</td>
<td>102.6 (11.5)</td>
<td>103.4 (15.6)</td>
</tr>
<tr>
<td>LOC ≥ 1, &lt;12 (N = 25)</td>
<td>91.6 (12.4)</td>
<td>93.7 (12.35)</td>
<td>94.0 (13.55)</td>
<td>93.1 (15.6)</td>
<td>93.0 (16.8)</td>
</tr>
<tr>
<td>LOC ≥ 12 (N = 20)</td>
<td>93.3 (15.2)</td>
<td>96.7 (11.8)</td>
<td>96.5 (17.4)</td>
<td>97.5 (16.57)</td>
<td>95.7 (18.1)</td>
</tr>
<tr>
<td>Total (N = 87)</td>
<td>94.9 (12.6)</td>
<td>96.1 (11.5)</td>
<td>97.7 (14.9)</td>
<td>98.7 (14.4)</td>
<td>98.7 (17.0)</td>
</tr>
</tbody>
</table>

Note. Values enclosed in parentheses represent standard deviation. LOC, loss of consciousness in hours; VIQ, Verbal Intelligence Quotient; WRAT-III, Wide Range Achievement Test—Third Edition; WMS, Wechsler Memory Scale; AI, Auditory Immediate memory; AD, Auditory Delayed memory; GM, General memory.
2.2. Measures

Measures used in prediction of memory scores were limited to the verbal domain in order to maximize estimation of premorbid ability, as single word-reading ability has been shown to be robust to cognitive decline following brain injury (Putnam, Ricker, Ross, & Kurtz, 1999), and verbal scales from the WAIS-III are more highly correlated with single word reading than either performance scales or Full Scale IQs. Intellectual functioning was measured with the Verbal Intelligence Quotient (VIQ) from the WAIS-III. Premorbid functioning was estimated with the Reading subtest of the WRAT-III (Wilkinson, 1993), as it has been shown to be an appropriate measure or premorbid functioning (Johnstone, Callahan, Kapila, & Bouman, 1996; Johnstone et al., 1995). Memory was measured with the Auditory Immediate, the Auditory Delayed, and the General memory indices from the WMS-III. Isolated visual memory measures from the WMSIII were not used due to their relatively low correlation with verbal measures in healthy individuals (The Psychological Corporation, 1997).

2.3. Procedures

Participants were administered the Wechsler Adult Intelligence Scale—Third Edition (WAIS-III, Wechsler, 1997a) and the Wechsler Memory Scale—Third Edition (WMS-III; Wechsler, 1997b) subtests, and the Wide Range Achievement Test—3rd Edition (WRAT-III, Wilkinson, 1993) reading subtest as part of a larger comprehensive neuropsychological evaluation administered by three psychometrists trained in standardized test administration. All psychometrists had at least a bachelor’s degree, and were formally trained and supervised in test administration over a period of at least 3 months prior to administering any measures independently. Two of the psychometrists had at least 2 years experience, while the third underwent training and began administering measures during the current project. Score discrepancies between general intellectual functioning and memory were calculated using three methodologies:

(1) The first approach is termed the Simple-difference method (The Psychological Corporation, 1997, p. 213). As detailed in the manual, the WMS-III score in question was subtracted from the WAIS-III VIQ score. Using the table provided in the manual, the difference was then analyzed for statistical rarity (i.e., was a reliable difference present), and if it was determined to be a reliable difference, population rarity was evaluated by examining how often the difference occurred in the standardization sample. For the current study, statistical significance for the difference scores was set at .05, and population rarity was defined as occurring in 10% or fewer individuals in the normative sample.

(2) The second approach is termed the Predicted-difference method (The Psychological Corporation, 1997, p. 214). In this approach, IQ scores were used to predict memory scores based on regression formulas derived from the standardization sample. The predicted memory score was then compared with the obtained memory score, and if the two scores were reliably different, population rarity was evaluated by examining how often the difference occurred in the standardization sample. For the current study, the
reliable difference for the predicted scores was set at .05, and population rarity was defined as occurring in 10% or fewer individuals in the general population.

(3) The third approach incorporated an estimate of premorbid functioning based on WRAT-III Reading score. Reading ability as measured by the WRAT-III has been shown to provide a reasonable estimate of premorbid functioning based on its high correlation with VIQ, and is commonly used in clinical practice (Johnstone et al., 1996). The WRAT-III Reading standard score was used to predict memory scores using a regression-based equation adapted from procedures used to identify learning disability discrepancies based on reliable differences (see Fig. 1; Reynolds, 1990). The critical value for determining reliable differences (i.e., $z_a$ in Fig. 1) was a traditional cut-off value of .05.

A difference was classified as clinically meaningful when the following criteria were met:

$$Y_{1} - Y_{2} > SD_{Y} z_{a} \sqrt{1 - r_{yx}^2}$$

where:

$Y_{1}$ is the patient’s memory index score

$SD_{Y}$ is standard deviation of $Y$

$z_{a}$ is the point on the normal curve corresponding to the relative frequency needed to indicate “impairment”

$r_{yx}$ is the square of the correlation between the aptitude and achievement measures

$\hat{Y}$ is the mean memory for all individuals with IQ = $X_{1}$

where

$$\hat{Y} = \left[ r_{yx} \left( \frac{X_{1} - \bar{X}}{SD_{Y}} \right) \right] SD_{X} + \bar{X}$$

where

$X_{1}$ = the patient’s VIQ score

$\bar{X}$ = the mean of the VIQ measure (i.e. for the WAIS-III = 100)

$SD_{X}$ = the standard deviation of the VIQ measure (i.e. for the WAIS-III = 15)

and

$r_{yx}$ = the correlation between $X$ and $Y$

Fig. 1. Formulas used to calculate reliable differences between estimated premorbid IQ scores and memory scores.
which results in using a difference of two standard deviations for $z_a$. For the equations provided in Figure 1, it was assumed that the correlation between the estimated VIQ and WMS-III index scores was .50 in a non-clinical sample. This is likely a high estimate of the correlation (Williams, 1997), however, smaller correlation estimates result in prohibitively large predicted score confidence intervals (Williams, 1997).

3. Results

Basic descriptive test data are reported in Table 1. As expected, means are slightly below the standardization sample ($M = 100$) secondary to the effects of TBI on all measured abilities. Impairment classification rates based on a level of a statistically reliable difference for the group data are shown in Figure 2. Group data for impairment classification rates representing differences that were rare in the population sample are shown in Figure 3. Overall, the Predicted-difference method tended to classify the fewest individuals as impaired based on a
Table 2
Percentage of overlap and Kappa statistics for agreement between three classification methods

<table>
<thead>
<tr>
<th>Memory index</th>
<th>Pred–Simp % overlap/Kappa</th>
<th>Pred–Read % overlap/Kappa</th>
<th>Simp–Read % overlap/Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical rarity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auditory Immediate</td>
<td>62.5/.731</td>
<td>43.7/.550</td>
<td>61.1/.711</td>
</tr>
<tr>
<td>Auditory Delayed</td>
<td>57.1/.688</td>
<td>69.2/.792</td>
<td>60.0/.710</td>
</tr>
<tr>
<td>General</td>
<td>54.5/.636</td>
<td>68.4/.770</td>
<td>65.2/.730</td>
</tr>
<tr>
<td>Population rarity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auditory Immediate</td>
<td>55.6/.690</td>
<td>50.0/.645</td>
<td>25.0/.365</td>
</tr>
<tr>
<td>Auditory Delayed</td>
<td>60.0/.726</td>
<td>50.0/.639</td>
<td>57.1/.709</td>
</tr>
<tr>
<td>General</td>
<td>61.5/.730</td>
<td>58.3/.707</td>
<td>45.4/.587</td>
</tr>
</tbody>
</table>

Note. Pred–Simp, Predicted-difference–Simple-difference method; Pred–Read, Predicted-difference–Premorbid-estimation method; Simp–Read, Simple-difference–Premorbid-estimation method; % overlap, percentage of agreement for specific cases between each method.

level of reliable difference, while the regression-based Premorbid method tended to classify the fewest individuals as impaired based on criteria for rarity in the population. Non-parametric Cochran $Q$ statistics for repeated measures were performed to compare the proportions of individuals classified as impaired. Results indicated the only statistically different classification rates occurred in the category for population rarity for Auditory Delayed memory with the Predicted-difference classifying significantly more individuals as impaired than the other two methods (Cochran’s $Q = 7.00$, $P < .05$).

In order to examine the degree of agreement between methods, the percentage of overlap was evaluated for each of the measures in pairwise comparisons (see Table 2). This was done by dividing the number of patients classified as impaired in both methods by the total number of students classified as impaired by the two methods separately (Bennett & Clarizio, 1988). Using the reliable difference criteria for decline in memory, rates of agreement ranged from a high of 69.2% for the Predicted-difference and Premorbid methods for Auditory Delayed memory, to a low of 43.7% for the Predicted-difference and Premorbid methods for Auditory Immediate memory. When using population rarity as the criterion, rates of agreement ranged from a high of 61.5% for the Predicted-difference and the Simple-difference methods for General memory, to a low of 25.0% for the Simple-difference and the Premorbid methods for Auditory Immediate memory. As a further measure of degree of agreement, Kappa statistics were calculated for each of the pairwise comparisons (see Table 2). Using the criteria outlined by Landis and Koch (1977), one of the comparisons is classified as “fair,” two are classified as “moderate,” and 15 are classified as “substantial.” None reached criteria (i.e., .80) for “almost perfect.”

Data were subsequently broken down and analyzed according to the injury severity information in order to examine potential systematic bias associated with more or less severe injuries. Consistent with expectations associated with sequella of TBI, individuals with moderate and severe injuries showed generally lower scores for VIQ and memory measures. Differences between individuals with mild, moderate, and severe injuries were greatest for Auditory Delayed and General memory indices (see Table 1), with the moderate injury showing
Table 3
Proportion of individuals classified as impaired based on statistical rarity in each injury severity category using three methods of classification

<table>
<thead>
<tr>
<th>WMS-III Scale</th>
<th>Injury severity category</th>
<th>LOC &lt; 1 h</th>
<th>LOC = 1–12 h</th>
<th>LOC &gt; 12 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple-difference</td>
<td>Auditory Immediate</td>
<td>.19</td>
<td>.16</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>Auditory Delayed</td>
<td>.09</td>
<td>.16</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>.21</td>
<td>.24</td>
<td>.25</td>
</tr>
<tr>
<td>Predicted-difference</td>
<td>Auditory Immediate</td>
<td>.12</td>
<td>.08</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>Auditory Delayed</td>
<td>.05</td>
<td>.20</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>.09</td>
<td>.24</td>
<td>.20</td>
</tr>
<tr>
<td>Premorbid method</td>
<td>Auditory Immediate</td>
<td>.12</td>
<td>.16</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>Auditory Delayed</td>
<td>.09</td>
<td>.20</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>.12</td>
<td>.32</td>
<td>.25</td>
</tr>
</tbody>
</table>

**Note.** LOC, loss of consciousness.

* Criterion for impairment = 1 S.D. below estimated premorbid functioning.

significantly lower scores than the mild injury group for both indices, $F(2, 84) = 3.69, P < .05$ and $F(2, 84) = 3.46, P < .05$, respectively. Consistent with previous reports concerning VIQ in individuals sustaining TBI, there were no statistically significant differences between groups for VIQ (Martin, Donders, & Thompson, 2000).

Once again, impairment classification rates were calculated based on statistically significant differences for the two methods involving the WMS-III standardization sample. In addition, in order to approximate the lower threshold inherent in the statistical rarity criterion with the premorbid methodology, a threshold of 1 S.D. difference between estimated premorbid functioning and obtained memory scores was used. See Table 3 for relative impairment classification rates based on the three different methods in three different TBI severity groups. In general, the Simple-difference method led to the highest impairment classification rates, with the Predicted-difference method tending to provide the lowest impairment classification rates. When statistical rarity was used as the criterion, Cochran’s $Q$ statistics indicated the only methods that showed differing classification rates were the Predicted-difference method and the Simple-difference method for individuals with no loss of consciousness on the General memory index ($Q = 7.00, P < .05$), with the Simple-difference method classifying significantly more individuals as impaired.

Impairment classification rates for differences that were unusual in the norming sample are shown in Table 4. As would be expected based on methodological design, fewer individuals were classified as impaired based on criteria of population-based rarity than were classified as impaired based on statistical significance across classification methodologies and injury severity categories. In this case, the regression-based Premorbid method (using the methodology illustrated in Fig. 1) tended to provide the lowest rates of classification across injury severity, with the Predicted-difference method providing the highest impairment classification rates.
Table 4

<table>
<thead>
<tr>
<th>WMS-III Scale</th>
<th>Injury severity category</th>
<th>LOC &lt; 1 h</th>
<th>LOC = 1–12 h</th>
<th>LOC &gt; 12 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple-difference</td>
<td>Auditory Immediate</td>
<td>.07</td>
<td>.04</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>Auditory Delayed</td>
<td>.02</td>
<td>.12</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>.09</td>
<td>.08</td>
<td>.15</td>
</tr>
<tr>
<td>Predicted-difference</td>
<td>Auditory Immediate</td>
<td>.09</td>
<td>.04</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>Auditory Delayed</td>
<td>.05</td>
<td>.20</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>.09</td>
<td>.20</td>
<td>.15</td>
</tr>
<tr>
<td>Premorbid method</td>
<td>Auditory Immediate</td>
<td>.05</td>
<td>.00</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>Auditory Delayed</td>
<td>.00</td>
<td>.12</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>.02</td>
<td>.12</td>
<td>.15</td>
</tr>
</tbody>
</table>

Note. LOC, loss of consciousness.

However, Cochran \( Q \) statistics revealed no significant differences in classification rates among any of the methodologies for any of the injury severity conditions. Rates of agreement and Kappa statistics were largely similar to those seen when the data were analyzed as a group.

4. Discussion

Contrary to hypotheses, general results did not indicate dramatic differences in overall impairment classification rates between three different classification methods for a clinical sample. Regression-based formulas based on the original WAIS-III–WMS-III norming sample provided impairment classification rates that were largely similar to those that involved estimates of premorbid functioning, with two notable exceptions. First, when differences that were rare based on the initial population sample (as opposed to statistically reliable differences) were considered, the Predicted-difference method classified individuals as impaired at a significantly higher rate than other methods for the Auditory Delayed memory index. Several factors may account for this finding. The regression-based methodology for meaningful differences involving premorbid estimates generally required difference scores of approximately 25 points on index scores, while the Predicted-difference method generally required differences of approximately 21 points. Thus, in cases with concordance between estimated and obtained VIQ, the Predicted-difference method has a greater likelihood of being significantly different. The added advantage to be gained by including premorbid estimates of functioning would be minimized in this scenario by the larger confidence intervals required from measures derived from different normative samples. The impact of this relatively large confidence interval associated with premorbid estimation of memory has been commented on by others (Williams,
It should also be noted that the correlations between estimated premorbid VIQ and memory measures were assumed to be at the high end of estimates (.50), suggesting that even fewer individuals would be classified as impaired if a more conservative estimate was used.

The second factor to consider in the results is the degree of decline in VIQ seen in the current sample. VIQ has been shown to be mildly sensitive to the effects of TBI (Johnstone et al., 1995; Skeel et al., 2000) though it may not be one of the most affected domains following TBI (Martin et al., 2000). While the estimated decline in the current sample was 5–8 points in the groups that had sustained moderate to severe TBI compared to the group that had sustained mild TBI, this may not have been sufficiently large to dramatically impact difference scores. A related factor concerns the utility of the comparisons across injury severity classification. As there were not significant differences in classification rates in any of the injury severity categories across methodologies, one method did not distinguish itself as more or less sensitive for a specific TBI population (e.g., premorbid estimation being differentially sensitive to moderate TBI). In clinical samples with populations that sustain larger declines in VIQ (e.g., Alzheimer’s disease, left-sided cerebrovascular accident, etc.) methodologies involving premorbid estimation would clearly provide increasingly accurate estimates of impairment.

Limitations of the current study include the degree of applicability of the current results to different settings. While every effort was made to maximize diversity, the setting and demographic characteristics of the surrounding community limited the variability of the population. Thus, it is possible that studies utilizing individuals with more diverse ethnic backgrounds may find different results. A second limitation is potential systematic bias related to background characteristics of the sample, such as SES, that may have artificially depressed or inflated IQ scores and/or reading levels. Since both measures were used to estimate expected memory performance, a systematic bias may have obscured differences that may have occurred in alternative populations.

Clinically, the current study clearly illustrates the advantages of using measures that have been co-normed. The confidence intervals for the co-normed measures allow for smaller differences between scores to be interpreted as significant, i.e., they allow for increased sensitivity to detect declines that are present. With regard to practical implications for interpretation, the current study suggests that it may be appropriate to use both estimates of premorbid functioning and Predicted-score methods when interpreting memory scores from the WMS-III. This is particularly relevant based upon the lack of perfect agreement seen between methods, and the fact that none of the methods demonstrated a distinct advantage over the others. Results suggest that despite the impact of TBI on VIQ, predicted scores from the WAIS-III–WMS-III manual guidelines tend to provide sensitivity at least equal to that of Premorbid-estimation methods. If that comparison does not prove to be significant, it is appropriate to examine memory scores in comparison to regression-based estimates of premorbid functioning to evaluate potential clinically significant decline. This second methodology is clearly most relevant when there is evidence of substantial decline in VIQ. In the event of differing classifications, clinicians are forced to rely on clinical judgment with regard to which methodology most accurately documents current functioning. It is possible that hybrid methods (Axelrod, Vanderploeg, & Schinka, 1999; Krull, Scott, & Scherer, 1995; Vanderploeg & Schinka, 1995) of premorbid estimation may provide for more accurate estimates of premorbid functioning in individual
cases, but these methods would continue to be affected by the large estimate range inherent in the relatively low correlation between estimated IQ and memory functioning.

A third important point is the illustration of the substantial differences in impairment classification rates seen when impairment is based on population rarity versus statistically reliable differences. When the Simple-difference method was used, less than one half of the individuals classified as showing statistically significant evidence of decline were classified as impaired when population rarity was used as the criterion. This difference was less pronounced for the Predicted-difference method. However, with premorbid estimation methods, when a criterion of 1 S.D. decline was used as evidence of decline, as few as one sixth of the sample originally classified as impaired continued to be classified as impaired when regression-based premorbid estimates were applied. As 1 S.D. difference has been advocated as suggestive of mild decline (Lezak, 1995), it is important for clinicians to recognize defining someone as impaired on the basis of a 1 S.D. is a tenuous position based on statistical analysis.

A final important note is the potential advantage of using reading-based measures of premorbid functioning that have been empirically linked with the WAIS-III and WMS-III. This would allow for direct substitution of VIQ estimated through reading ability in order to predict memory functioning directly from reading ability. This would have the potential to optimize confidence intervals necessary to detect differences between estimated premorbid functioning and current functioning that were rare in the initial norming sample.

Future research may include replicating the current findings in alternative settings in order to expand generalizability. The current sample was limited to individuals with TBI. In addition, it may be useful to examine individuals who have different disorders, such as Alzheimer’s disease or cerebrovascular disease, in order to explore the pattern of deficits and determination of impairment in alternative populations. Finally, future research may also examine the degree to which alternative estimations of premorbid intelligence may affect the sensitivity of the various methodologies to detect impairment.

References


