Journal of the International Neuropsychological Society (2001), **7**, 875–880.

The influence of IQ Stratification on WAIS–III/WMS–III FSIQ–General Memory Index discrepancy base-rates in the standardization sample

KEITH A. HAWKINS[1] AND DAVID S. TULSKY[2]

[1]Yale University School of Medicine, New Haven, Connecticut
[2]Kessler Medical Rehabilitation Research and Education Corporation, West Orange, New Jersey

(RECEIVED February 8, 2001; REVISED April 16, 2001; ACCEPTED May 21, 2001)

Abstract

Since memory performance expectations may be IQ-based, unidirectional base rate data for IQ-Memory Score discrepancies are provided in the *WAIS–III/WMS–III Technical Manual*. The utility of these data partially rests on the assumption that discrepancy base rates do not vary across ability levels. FSIQ stratified base rate data generated from the standardization sample, however, demonstrate substantial variability across the IQ spectrum. A superiority of memory score over FSIQ is typical at lower IQ levels, whereas the converse is true at higher IQ levels. These data indicate that the use of IQ–memory score unstratified "simple difference" tables could lead to erroneous conclusions for clients with low or high IQ. IQ stratified standardization base rate data are provided as a complement to the "predicted difference" method detailed in the *Technical Manual*. (*JINS*, 2001, *7*, 875–880.)

Keywords: WAIS–III/WMS–III, Intelligence, Memory, Discrepancy, Base-rates

INTRODUCTION

One method of examining for memory decline has been to compare memory and FSIQ data. Discrepancies of the order of 15 points or 20 points have been explored as supporting the possibility of a loss in memory capacity or the presence of a condition characterized by memory impairment, subject to the usual caveat that findings be interpreted in the light of history, suspected presence of a relevant injury or disease process, other cognitive data, and so forth (Bornstein et al., 1989; Butters & Cermack, 1980).

The conorming of the WMS–III with the WAIS–III places IQ–memory data discrepancy analysis on firmer footing, since score comparisons previously incorporated error variance due to the derivation of scores from differing normative bases. Tables providing base-rate data for unidirectional differences (i.e., for differences resulting from the subtraction of memory score from IQ) are provided in the *Technical Manual* (Psychological Corporation, 1997). Although a major advance, the utility of these tables across the full range of examinees rests in part upon the assumption that IQ–memory discrepancy base rates do not differ across the ability spectrum. There is reason to doubt that this is the case, since difference scores between two variables will reflect a regression to the mean effect inversely proportional to the correlation between them (Horowitz, 1974). When one score is high, the contrasted score is likely to be lower; when one is low, the other is likely to be higher. Highly correlated variables display smaller regression to the mean effects than more weakly related variables. Since FSIQ and GMI correlate only moderately in the standardization sample (.60; Psychological Corporation, 1997) large effects of IQ level upon intellectual–memory discrepancy base rates are to be expected: high IQ subjects will show lower memory scores, and low IQ subjects higher memory scores.

One way to understand the relationship of variable correlation to discrepancy magnitude is to consider a situation where the variables do not correlate ($r = 0$). In this case, the best prediction of the mean of a second score (e.g., memory) given knowledge of the first (e.g., IQ), is 100, no matter how bright (or dull) the sample is. A sample with a mean IQ of 125 will show very large IQ–memory discrepancies (averaging 25 points or so), and a group with a mean IQ of 75 will show a similarly large discrepancy but in the oppo-

Reprint requests to: Keith A. Hawkins, Yale University School of Medicine, CMHC, Room 530, 34 Park Street, New Haven, CT 06519. E-mail: keith.hawkins@yale.edu

https://doi.org/10.1017/S1355617701777119 Published online by Cambridge University Press

site direction, with memory considerably higher. The IQ–memory discrepancy magnitude and directionality displayed within the standardization data should represent an intermediate position between this example (of no correlation) and the situation that would emerge if IQ and memory scores correlated perfectly, in which case all discrepancies will be zero. The purpose of this study is to explore the extent of IQ–GM discrepancy base-rate variability across IQ levels, and to generate IQ-stratified base-rate data to complement those provided in the *Technical Manual*.

METHODS

Sample

The study sample consists of the WMS–III standardization sample (weighted $N = 1250$ adults) containing individuals 16 to 89 years of age administered both the WAIS–III and WMS–III. Inclusion criteria and demographics are detailed in the Technical Manual for these instruments (Psychological Corporation, 1997), with further elaboration in Tulsky and Ledbetter (2000).

RESULTS

Base rates for FSIQ–General Memory Index discrepancies varied considerably with level of FSIQ.

Discrepancy Directionality

There is virtually an even split between cases exhibiting a superiority of IQ and cases exhibiting a superiority of GMI within the standardization sub-sample with FSIQs between 90 and 109. However, FSIQ exceeded GMI in just 16.1% of cases in the sub-sample of individuals with FSIQs lower than 80. In striking contrast, at IQ levels of 120+ the trend is reversed, with FSIQ exceeding GMI in 86.6% of cases. This shift in dominance from the lower IQ levels, where memory is superior, to the higher, where IQ dominates, is remarkably regular across the IQ bands, as seen in Figures 1 and 2.

Discrepancy Magnitude

Since there is greater latitude for differences to emerge at higher score levels, larger IQ–GMI discrepancies might be expected at higher ability levels. Larger discrepancies *are* somewhat more common with higher levels of FSIQ, but this trend is not as pronounced as might be expected. The standard deviations for the discrepancy distributions differ only modestly between the lowest and highest FSIQ groups (10.4 cf. 12.1), reflecting distribution curves that are similar in breadth (see Figure 3).

Discrepancy Base Rates

Whereas a superiority of FSIQ over GMI of 15 points is seen in only 1% of IQ cases below 80, and in less than 4% of cases with FSIQ less than 90, over 33% of standardization cases with an IQ exceeding 119 exhibit such a discrepancy. Since the unusualness of a given FSIQ-GMI discrepancy varies dramatically with level of IQ, we have generated base rate data for clinical reference purposes. Table 1 presents directional base rate data from the standardization sample (that is, the frequency data are always for FSIQ minus GMI).

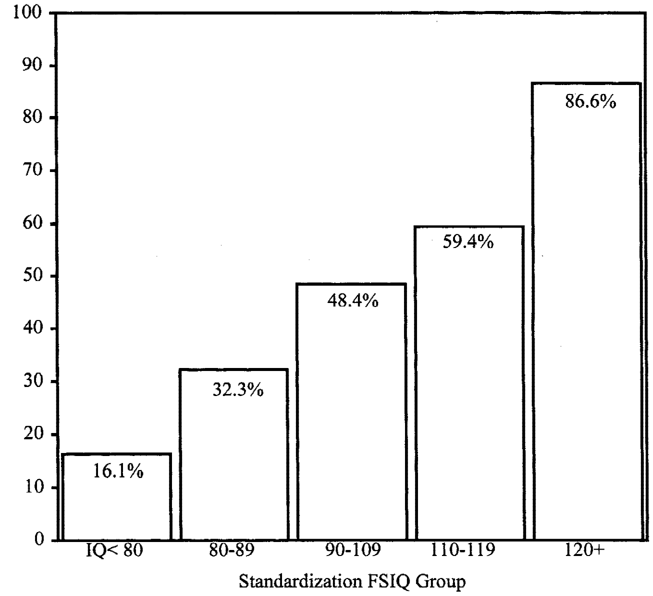

Fig. 1. Percentage of sample with FSIQ > GMI by FSIQ level.

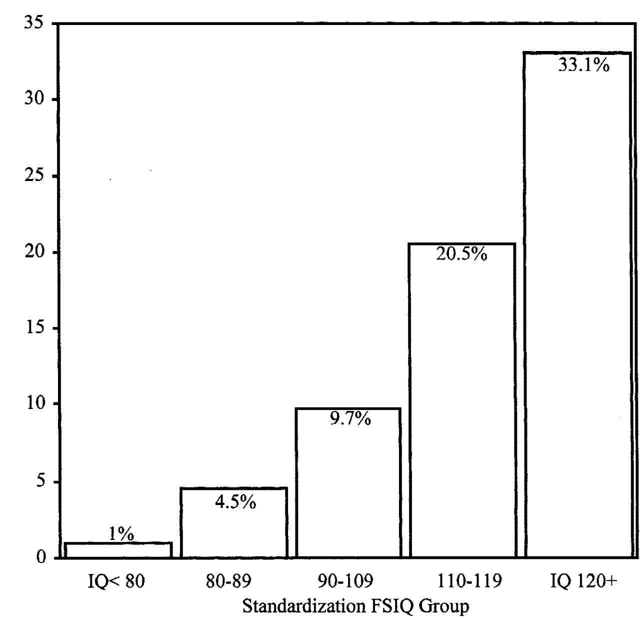

Fig. 2. Percentage of sample with FSIQ > GMI by 15+ by FSIQ level.

https://doi.org/10.1017/S1355617701777119 Published online by Cambridge University Press

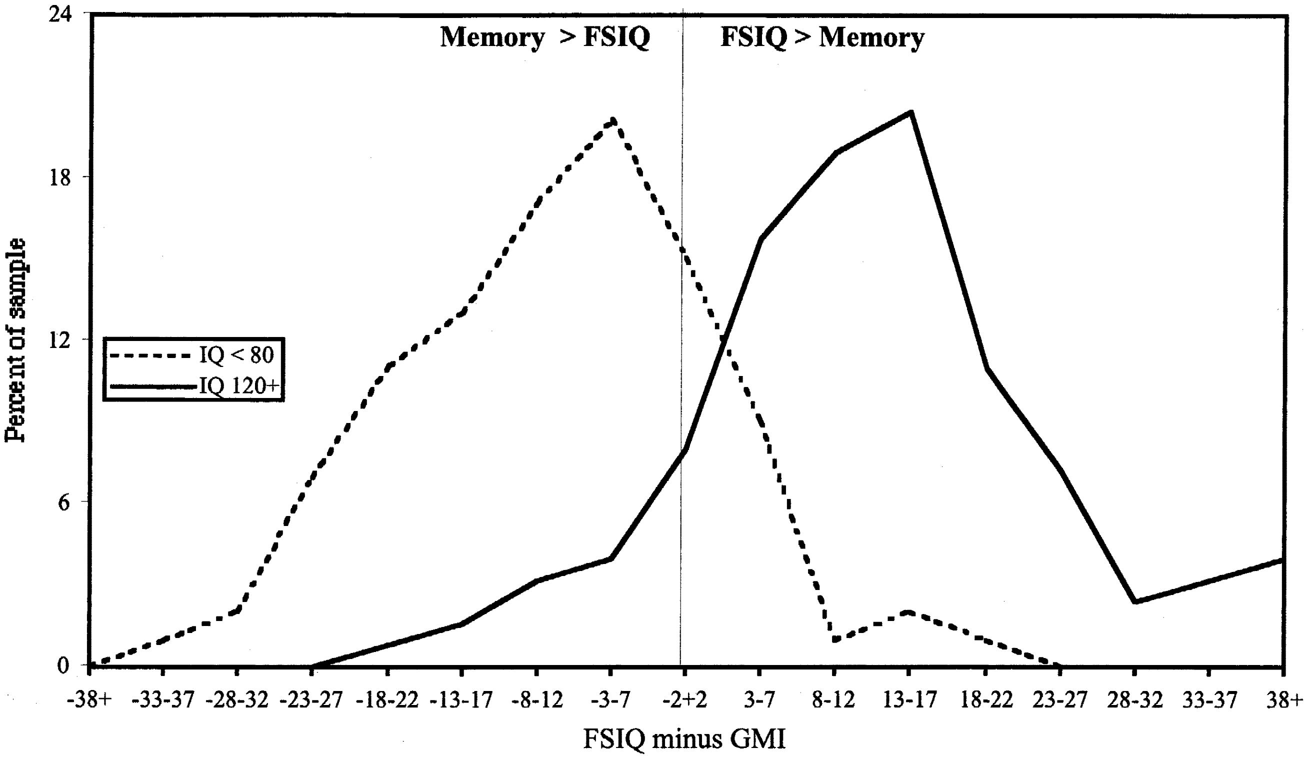

Fig. 3. Distributions of discrepancy scores for standardization sample group IQ < 80 and IQ group 120+.

DISCUSSION

The existence of substantial IQ-level dependent variability in the base-rates for IQ–GMI discrepancies has profound implications for the interpretation of IQ–Memory Index discrepancy data. In one of the following hypothetical cases, the magnitude of the superiority of IQ over GMI is both statistically significant and uncommon. In the other case, the difference is neither significant nor rare.

Case 1 is a 35-year-old with memory complaints who generates a FSIQ of 129 and a General Memory Index of 108. Case 2 is a 35-year-old with memory complaints who generates a FSIQ of 79, and a General Memory Index (GMI) of 73. Since FSIQ exceeds memory by 21 points in the 1st case, but only 6 in the 2nd, it might seem obvious that Case 1 rather than Case 2 features the statistically significant and uncommon, and therefore potentially meaningful, discrepancy.

In fact, due to the impact of IQ upon the base rates, Case 2 features the significant and unusual discrepancy. Using the predicted difference method, the GMI predicted by a FSIQ of 79 is 87 (*Technical Manual*, p. 264). The difference between the actual GMI score (73) and this prediction is 14 points, which is statistically significant at .05 for subjects of this age (Table B.4 of the *Technical Manual*; Psychological Corporation, 1997). In addition, consistent with the finding that at lower IQ levels memory scores typically exceed IQ, 94% of the Standardization Sample with IQ less than 80 exhibit a *smaller* IQ minus GM difference; just 6% have an IQ superiority of 6 points or greater (Table 1, this paper).

In contrast, the 9-point difference between Case 1's predicted (117) and obtained (108) GMI scores is not statistically significant. Although the difference between the actual IQ and GMI scores of 21 points seems large, it is exceeded by 17.3% of the FSIQ = 120+ segment of the standardization sample.

IQ–Memory Discrepancy Analysis: Clinical Considerations

IQ level and differential risk for false positive or negative errors

The preceding case studies demonstrate a differential risk for different types of diagnostic error across IQ strata. With lower IQ subjects there is a heightened false negative risk: IQ–GMI discrepancies are likely to be interpreted as not providing evidence for memory decline, when in fact memory decline has occurred. A higher GMI (than FSIQ) is the norm; therefore, counterintuitive though it seems, a small discrepancy in favor of IQ may speak to a relative weakness in memory (or at least, an imbalance between IQ and memory relative to the normal picture in similarly intellectually endowed individuals). Neuropsychologically, this may be important to note when assessing for early Alzheimer's disease, and in traumatic brain injury, where patients with a lower than average premorbid IQ appear to be dispropor-

https://doi.org/10.1017/S1355617701777119 Published online by Cambridge University Press

Table 1. FSIQ–GMI discrepancy standardization sample base rates stratified by FSIQ

FSIQ minus GMI	Cumulative percentages stratified by FSIQ							FSIQ minus GMI
	FSIQ < 80	80–89	90–99	100–109	110–119	120–129	FSIQ 130+	
−2	74							−2
−1	78							−1
0	84							0
1								1
2	87	74						2
3	88	77						3
4		78						4
5	94	79	73					5
6	95	84	76					6
7	96	87	80					7
8	97	87	81					8
9		88	84	73				9
10		91	85	77				10
11		92	87	81	75			11
12		94	89	84	76			12
13	99	95	90	85	78			13
14		96	93	88	80			14
15			94	89	83	73		15
16			94	91	85	77		16
17		96	95	92	86	78		17
18			96	93	89	78		18
19	100	98	96	95		80		19
20			96	96		87		20
21		99	97	97	90	88		21
22			98	97				22
23				98	92	90		23
24			98		94	92		24
25			98	98	95			25
26		99	99		96		70	26
27			99		97		85	27
28		100			97	93		28
29								29
30					97			30
31								31
32					99	94	90	32
33			100	100				33
34						96		34
35								35
36					100	97		36
37								37
38						98		38
39								39
40						99		40
41						100		41
42							95	42
43								43
44								44
45								45
46								46
47							100	47

tionately represented (Putnam & Adams, 1992; Rimel et al., 1981).

Conversely, given that IQ increasingly outstrips GMI as FSIQ rises above the average range, so too does the risk for false positive decisions. Clinicians unaware that one-third of subjects with IQ exceeding 119 exhibit a discrepancy of 15 or higher are at risk of interpreting such data as indicating memory decline.

Implications of insult-associated IQ decline

Inferences based upon a discrepancy between a memory score and IQ could be subject to error due to the probability that IQ has been depressed along with memory by the insult at issue. A drop in IQ and memory score in roughly equal measure will obviously obscure memory decline. When there is reason to believe that intellectual capacity has also suffered, the lack of a discrepancy between memory and IQ obviously should not be employed to infer that there has *not* been a loss in memory capacity.

But what if memory loss is suspected based on the probable presence of a condition known to have a differentially greater impact upon memory? How does a coexisting, though lesser, decline in IQ play out with regard to discrepancy analysis?

This is less of a problem than it first appears, as the effect of the IQ decline is countered by two factors acting together: the greater impact upon memory of the condition, combined with the increasing rarity of any given discrepancy as the IQ strata are descended. This is best explained via some additional hypothetical examples (not to be confused with the cases discussed earlier), where we assume divine knowledge of the premorbid IQ for each case. For simplicity, we will further assume that each referral involves suspected early Alzheimer's disease (AD), a condition for which memory impairment is prototypical and a common reason for neuropsychological referral. We will look at several examples spanning the IQ spectrum.

In Case 1, the premorbid IQ of 125 has declined to 118 at the time of diagnostic testing. The GMI at testing is 92. The difference between the diagnostic testing FSIQ (118) and GMI (92) is 26 points, a discrepancy seen in less than 5% of standardization sample cases in the FSIQ 110–119 stratum (Table 1). The premorbid IQ *versus* current GMI discrepancy (125 *vs.* 92) of 33 is larger, but of no greater rarity when assessed against the base rates for the applicable premorbid IQ stratum (IQ 120–129).

In Case 2, the premorbid IQ has dropped from 112 to 98, and the diagnostic testing GMI is 82. The discrepancy in hand of 16 is seen in no more than 6% of IQ 90–99 cases, but in 15% of high average cases. The large 30-point discrepancy between premorbid IQ and current GMI is of greater rarity, but when matched against the premorbid (high average) IQ stratum base-rates the difference (cumulative percentage 97 *vs.* 94; Table 1) is far from critical.

In Case 3, the decline is from a premorbid IQ of 97 to 88, and the diagnostic testing GMI is 75. The current discrepancy of 13 is exceeded by no more than 5% of low average IQ standardization cases, a base rate only marginally greater than that for the premorbid IQ–current GMI discrepancy of 22 within the average IQ group.

In Case 4, the FSIQ drops from 85 to 78, and the GMI is 70. The current 8-point discrepancy is equaled or exceeded by 3% of borderline IQ cases, virtually identical to the low average IQ base rate for the 15-point premorbid IQ–current GMI discrepancy.

In each of these examples the decline in IQ accompanying the onset of the condition has not resulted in substantially greater danger of a false negative clinical decision. Each, however, involves a decline in IQ from a higher to lower IQ stratum. What if the decline occurs within a stratum? This constitutes a threat, but not a great one, since any appreciable IQ decline (e.g., 6 points or greater) will typically result in the application of base rates from a lower stratum being applied. In generating FSIQ–GMI base rates we divided the average IQ standardization sample into two (90–99; 100–109), and the superior and above sample into 120–129 and 130+ groups, to lessen the frequency of within-strata IQ declines, since such declines will lessen the sensitivity of IQ-GM discrepancy to memory impairment. In any event, the clinician should view discrepancy analysis as just one source of inference, and should interpret the data in light of the broader neuropsychological exam, referral information, and personal history.

A critical implication of these base rate data is that, with regard to IQ–memory data, the unstratified simple differences method (Psychological Corporation, 1997) of determining the rarity of a given discrepancy and its statistical significance will be misleading much of the time. The predicted differences method offers greater protection against errors arising from regression effects (Horowitz, 1974; Reynolds, 1985) such as the vast differences in discrepancy direction that exist for IQ–Memory data across the IQ strata. The *Technical Manual* and scoring software (Psychological Corporation, 1997) offer predicted differences data between FSIQ, VIQ, or PIQ and WMS–III index scores. For the purposes of this paper we have opted to focus on FSIQ–GMI discrepancy base rates to complement these *Technical Manual* data, since the FSIQ and GMI constitute the largest, and likely most robust, of composite scores for their respective domains. A large number of other WAIS–III/WMS–III IQ-index and index–index contrasts are of course possible. Most such across-test comparisons manifest a similar influence of IQ on discrepancy base-rate direction, and to that end the FSIQ–GMI contrast serves as an exemplar for other pairings. Comprehensive WAIS–III Index–WMS–III Index base-rate data are in preparation and will be presented in the context of a larger discussion of discrepancy analysis as a method of neuropsychological inference (Hawkins & Tulsky, in press).

Although the provision of IQ stratified base rates enhances the utility of IQ–GMI discrepancy analysis, we wish to stress that the *validity of clinical interpretations* made on the basis of discrepancy infrequency is a more complex matter. We concur with Bornstein et al.'s (1989) warning that IQ–Memory Index discrepancies should not be used in isolation to identify memory deficits. In addition to the challenges posed by the nuances of memory and its disorders, we recognize that the diagnostic utility of WAIS–III/WMS–III discrepancy analysis (both within and across the tests) largely awaits determination. Preliminary clinical data show encouraging group discrepancies for several common conditions (Hawkins, 1998; Psychological Corporation, 1997), but sensitivity, specificity, and other relevant data such as diagnosis specific odds ratios (Bieliauskas et al., 1997; Ivnik et al., 2000) are largely lacking. At least in older subjects some composite scores may be less stable over time than orthodox reliability data indicate (Ivnik et al., 1995, 2000), a finding that possibly complicates these issues.

Despite these concerns, clinicians will continue to examine the relationship of intellectual to memory data in their neuropsychological analysis of memory complaints. The provision of IQ stratified FSIQ–GMI base rates leaves them better equipped to do so.

ACKNOWLEDGMENTS

Data presented in this article are drawn from the *Technical Manual of the Wechsler Adult Intelligence Scale/Wechsler Memory Scale–Third Edition*, Copyright 1997 by the Psychological Corporation or were calculated by subtracting GMI scores from FSIQ scores for subjects making up the standardization sample for these instruments. Adapted and reproduced by permission. All rights reserved. Portions of this paper were presented at the American Psychological Association Annual Convention, Washington DC 2000, and at the National Academy of Neuropsychology Annual Convention, Orlando, FL, 2000. We wish to thank Robert Ivnik, Roger Graves, Robert Knight and the *JINS* reviewers for their helpful comments.

REFERENCES

Bieliauskas, L.A., Fastenau, P.S., Lacy, M.A., & Roper, B.L. (1997). Use of the odds ratio to translate neuropsychological test scores into real-world outcomes: From statistical significance to clinical significance. *Journal of Clinical and Experimental Neuropsychology*, *19*, 889–896.

Bornstein, R.A., Chelune G.J., & Prifitera, A. (1989). IQ–memory discrepancies in normal and clinical samples. *Psychological Assessment*, *1*, 203–206.

Butters, N. & Cermack, L.S. (1980). *Alcohol Korsakoff's syndrome*. New York: Academic Press.

Hawkins, K.A. (1998). Indicators of brain dysfunction derived from graphic representations of the WAIS–III/WMS–III Technical Manual clinical samples data: A preliminary approach to clinical utility. *Clinical Neuropsychologist, 12*, 535–551.

Hawkins, K.A. & Tulsky, David S. (in press). Discrepancy analysis with the WAIS–III/WMS–III. In D. Tulsky & D. Saklofske (Eds.), *New approaches to clinical interpretation using the WAIS–III and WMS III*. New York: Academic Press.

Horowitz, L.M. (1974). *Elements of statistics for psychology and education*. New York: McGraw-Hill.

Ivnik, R.J., Smith, G.E., Malec, J.F., Petersen R.C., & Tangalos, E.G. (1995). Long-term stability and intercorrelations of cognitive abilities in older persons. *Psychological Assessment*, *7*, 155–161.

Ivnik, R.J., Smith, G.E., Petersen R.C., Boeve, B.F., Kokmen, E., & Tangalos, E.G. (2000). Diagnostic accuracy of four approaches to interpreting neuropsychological test data. *Neuropsychology*, *14*, 163–177.

Psychological Corporation. (1997). *WAIS–III WMS–III technical manual*. San Antonio, TX: The Psychological Corporation.

Putnam, S.H. & Adams, K.M. (1992). Regression-based prediction of long-term outcome following multidisciplinary rehabilitation of traumatic brain injury. *Clinical Neuropsychologist*, *6*, 383–405.

Reynolds, C.R. (1985). Critical measurement issues in learning disabilities. *Journal of Special Education*, *18*, 451–476.

Rimel, R.W., Giordani, B., Barth, J.T., Boll, T.J., & Jane, J.J. (1981). Disability caused by minor head injury. *Neurosurgery*, *9*, 221–228.

Tulsky, D.S. & Ledbetter, M.F. (2000). Updating to the WAIS–III and WMS–III: Considerations for research and clinical practice. *Psychological Assessment*, *12*, 253–262.