

Review of the WAIS-IV

The measurement of low IQ with the WAIS-IV:  
A critical review

By

Simon Whitaker

Consultant Clinical Psychologist/Senior Visiting Research Fellow

The Learning Disability Research Unit

Room HW 2/08

University of Huddersfield

Queen's Gate

Huddersfield

HD1 3DH

UK

Tel 01484 473612

[s.whitaker@hud.ac.uk](mailto:s.whitaker@hud.ac.uk)

in Clinical Psychology Forum, 45-48, 231, March 2012

## **Summary**

The WAIS-IV is reviewed for its use in the low IQ range. It has a floor that will elevate low score, does not give accurate percentile ranking and may well systematically measure higher than the WISC-IV.

## **Introduction**

The Wechsler Adult Intelligence Scale - Fourth Edition (WAIS-IV UK Wechsler 2008, 2010) is the latest standardisation of the widely used assessment of adult intelligence. It replaces the WAIS-III published in 1998 (Wechsler 1998) and is likely to be in common use for the next 10 to 12 years.

In many respects the WAIS-IV is an improvement on its predecessors. The Verbal and Performance IQs, which, together with Full Scale IQ, have been the basic derived scores on the WAIS since the Wechsler-Bellevue (Wechsler 1939) are no longer given. They are replaced by four index scores: Verbal Comprehension (VCI), Perceptual Reasoning (PRI), Working Memory (WMI) and Processing Speed (PSI). These index scores are in line with recent factor analytic studies of the WAIS-III (Taub, et al 2004) and the WISC-IV

(Wechsler 2008) and, according to Wechsler (2008), give the test better clinical utility. Although equivalent index scores could be derived from the WAIS-III, WMI and PSI could only be calculated if two additional subtests were given as well as the 11 core subtests required to derive a FS IQ. The WAIS-IV produces all four index scores and the Full Scale IQ with just ten subtests.

The core subtests have also changed so that the ones that are used now are designed to more specifically assess the abilities in the four index scales and to be quicker to administer. Other improvements include changing the artwork, and updating the record form, making scoring more convenient for the examiner.

The evidence for the reliability and validity of the WAIS-IV as outlined in the Technical and Interpretive Manual (Wechsler 2008) seems excellent. The main reliability figure based on the standardisation sample for Full Scale IQ was .98, which shows very good internal consistency. In addition, the test re-test reliability of .96 was found for Full Scale IQ when 298 were given the assessment twice with a mean interval of 22 days.

Considerable evidence is also presented with regard to the validity of the test. The test content is based on many years of experience and research in developing Wechsler assessments and expert opinion. Factor analytic studies (Wechsler 2008) demonstrate that the assessment is a good measure of general intelligence or *g* and of the cognitive abilities measured in the index scales. The test also correlates highly with other tests of general intelligence: for example, the WAIS-III ( $r = .94$ ) and the WISC-IV ( $r = .91$ ).

Finally, the test now comes in a handy rucksack rather than a briefcase as has been the case in its predecessors, which for the current author makes transporting it much more convenient.

It therefore seems that for adults of average intellectual ability the WAIS-IV is valid and reliable, has greater clinical utility, is easier to administer and more appropriate for today than its predecessor. However, as one of the main uses of intelligence tests is the assessment of adults with low intellectual abilities, it needs to be demonstrated that it is also appropriate for this group. It has been questioned whether the WAIS-III and its child equivalent, the WISC-IV, are sufficiently accurate in measuring low IQ. Whitaker (2005; 2008a) has suggested that some of the instructions are too complex for people at the low

IQ level to be able to understand. The subtest of main concern was Letter Number Sequencing, where the WAIS-III instructions were 80 words long and required the client to understand the concept "in alphabetical order".

Whitaker (2005, 2010) and Whitaker and Wood (2008) have suggested that allocating scaled scores of one to very low raw scores and raw scores of zero will artificially elevate IQ and index scores for some individuals who score in the low IQ range.

Whitaker (2005, 2008a) has pointed out that percentile ranking on percentiles less than one on the WISC-IV and WAIS-III are too low. This is because intellectual ability is not normally distributed at the very low IQ levels, with the prevalence of severe to profound intellectual disability (IQ<50) being found to be about 0.4% of the population as a whole (Abramowicz and Richardson 1975, Roeleveld et al 1997), rather than the <0.1% that would be predicted if IQ was normally distributed at these IQ levels.

Whitaker (2008b, 2010) noted that the reliability figures reported in the manuals are those for people with an average level of intellectual ability and suggests they may not be the same for people with low intellectual ability. He

has also noted that on the 90% and 95% confidence interval for IQ and index scores given in the administrative manuals are based on only one reliability score, in the main that for internal consistency, which does not take into account all the error affecting the test and so does not give a true estimate of the accuracy of the tests.

Whitaker (2003, 2008a, 2010) has pointed out that the agreement between IQ tests at the low levels may not be as good as it is in the average range. This was tested in the case of the WISC-IV (UK) and WAIS-III (UK) by Gordon et al (2010) who gave both tests to 16-year-olds in special education. It was found that the WISC-IV Full Scale IQ was always lower than the WAIS-III's with a mean difference of just less than 12 points.

### **The WAIS-IV and the measurement of low IQ**

The extent to which the above concerns with regard to the measurement of low IQ in early versions of the Wechsler tests are still applicable in the WAIS-IV are considered below.

## **Instructions**

The issue with instructions being too complex for people with low intellectual ability seems to be reduced. Instructions are given in manageable chunks broken up, with the client being shown examples of what they are required to do. The instructions to Letter Number Sequencing, which were 80 words long in the WAIS-III are now split up between demonstrations, the longest sequence being 33 words. However, the instructions still use the phrase "in alphabetical order" which is a concept that one could not automatically expect clients with low IQs to be familiar with.

## **Reliability**

As part of the validity of the WAIS-IV it was given to various special groups including 73 people with a mild intellectual disability and 31 people with a moderate intellectual disability. The reliability coefficients of the internal consistency are calculated and given in the manual for those sub-tests where this could be done. It is clear from this data that the internal consistency of the test at the low level is comparable with the average IQ level. However, no test re-test reliability is reported at the low level. Also, as with all other versions of the Wechsler tests, the 95% confidence interval is based on only one reliability score which ignores a lot of the error in the measurement of IQ.

## **Floor Effect**

Scaled scores of one are still allocated to very low raw scores and raw scores of zero, which will result in a floor effect on scaled score, which may well elevate FS IQ and index scores. As with other Wechsler tests it is indicated in the manual that a FS IQ or relevant index scores should not be derived if a client gets a raw score of zero on two of the three subtests measuring VCI or PRI or both subtests measuring WMI or PSI. However, it has been demonstrated (Whitaker 2010) that there is still ample scope for a floor effect by allocating scaled scores of one to low but none zero raw scores.

## **Percentile ranking**

As with previous Wechsler test the percentile ranking given in the manual are based on an assumption that IQ continues to be normally distributed below the range of IQs covered by the standardisation sample. It therefore allocates a percentile ranking of 0.4% for a FS IQ of 60 and of <0.1 to all IQs of 50 or less. As it was noted above the evidence suggests that the proportion of the population with IQs of 50 or less is about 0.4%, so the percentile rankings given in the manuals are a gross underestimate of the true level in the population.



### **Agreement with the WISC-IV**

As part of its validation the WAIS-IV was compared with the WISC-IV using people of average intellectual ability. The mean FS IQ on the WAIS-IV was 102.5 and that on the WISC-IV 103.7, a difference of 1.2 points with a correlation of .91, showing that the two tests are comparable in the average IQ range. However, no data is provided on how well the two assessments agree in the low range. As Whitaker and Wood (2008) and Whitaker (2008b) found evidence for a significant difference between the WAIS-III and the WISC-III and WISC-IV by comparing the criteria for a 16-year-old to obtain a scaled score of two on the common core subtests, this analysis was repeated to compare the WAIS-IV (Wechsler 2010) with the WISC-IV (Wechsler 2002). This indicated that it is harder to get a scaled score of two on the WISC-IV than it is on the WAIS-IV, which would mean that the WAIS-IV will probably score systematically higher than the WISC-IV at low IQ levels. To give some examples of common core subtests:

On Coding, which is virtually the same sub-test on both assessments, a 16-year-old would require a raw score of 35 on the WISC-IV and 23 on the WAIS-IV for a scaled score of two. With Symbol Search, which has a slightly different

response format on the two tests, but is very much the same subtest, the WISC-IV requires a raw score of 16 and the WAIS-IV of 10.

With Block Design the raw score 13 is required on the WISC-IV and eight on the WAIS-IV. This would be achieved on the WISC-IV if the 16-year-old completed the first five items correctly, four of which were 4 block items and one is 2 block item. On the WAIS-IV the scaled score of two would be gained if the adolescent completed only four items, two of which would be 2 blocks and two 4 blocks.

On the Vocabulary sub-test the raw score necessary for a 16-year-old to gain a scaled score of two is 23 on the WISC-IV and only one on the WAIS-IV. On the WISC-IV the client would be expected to give definitions to such words as "Thief", "Bicycle", "Alphabet" and "Leave". On the WAIS-IV they would only be required to correctly define "Book".

Although what is needed is an empirical study in which both tests are given to 16-year-olds in the low IQ range, this preliminary analysis does suggest that it is probably harder to gain a scaled score of two on the WISC-IV than it is on

the WAIS-IV and so it is likely that in the low IQ level the WISC-IV will systematically score lower than the WAIS-IV.

## **Discussion**

In many respects the WAIS-IV is an improvement on the WAIS-III, being easier, less time-consuming to administer and probably more user-friendly for the clients. However, there are still some major concerns with regard to its use on people with low intellectual abilities. It is not clear how stable the test is in the low range. It remains uncertain as to the degree to which Full Scale IQ and index scores are artificially increased by a floor effect. The percentiles cited in the manual below one percent would appear to be incorrect and the FS IQ and index score may well be systematically higher than their equivalents on the WISC-IV.

## Reference

Abramowicz, H.K. & Richardson, S.A. (1975). Epidemiology of severe mental retardation in children: Community studies. *American Journal of Mental Deficiency, 80*, 18-39.

Gordon, S., Duff, S., Davidson, T. and Whitaker S. (2010). Comparison of the WAIS-III and WISC-IV in 16-Year-Old Special Education Students. *Journal of Applied Research in Intellectual Disabilities, 23*, 197-200.

Roeleveld, N., Zielhuis, G.A. & Gabreels, F. (1997). The prevalence of mental retardation: A critical review of recent literature. *Developmental Medicine and Child Neurology, 39*, 125-132.

Taub, G.E., McGrew, K.S. & Witta, E.L. (2004). A confirmatory analysis of the factor structure and cross-age invariance of the Wechsler Adult Intelligence Scale-Third Edition. *Psychological Assessment, 16*, 85-89.

Wechsler, D. (1939). *Wechsler-Bellevue Intelligence Scale*. New York: The Psychology Corporation.

Wechsler, D. (1992). *Wechsler Intelligence Scale for Children - Third UK Edition: Administrative and Scoring manual*. London: The Psychology Corporation.

Wechsler, D. (1998). *Wechsler Adult Intelligence Scale - Third UK Edition: Administrative and Scoring manual*. London: The Psychology Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children - Fourth Edition: Technical and Interactive Manual*. San Antonio, TX: The Psychology Corporation.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale- Fourth Edition: Technical and Interactive Manual*. San Antonio, TX: The Psychology Corporation.

Wechsler, D. (2010). *Wechsler Adult Intelligence Scale - Fourth UK Edition: Administrative and Scoring Manual*. London: The Psychology Corporation.

Whitaker, S. (2003). Should we abandon the concept of mild learning disabilities? *Clinical Psychology, 29*, 16-19.

Whitaker, S. (2005). The use of the WISC-III and the WAIS-III with people with a learning disability: Three concerns. *Clinical Psychology, 50*, 37-40.

Whitaker, S. (2008a). WISC-IV and low IQ: review and comparison with the WAIS-III. *Educational Psychology in Practice, 24*, 129-137.

Whitaker, S. (2008b). The stability of IQ in people with low intellectual ability: An analysis of the literature. *Intellectual and Developmental Disabilities, 46*, 120-128.

Whitaker, S. (2010). Error in the estimation of intellectual ability in the low range using the WISC-IV and WAIS-III. *Personality and Individual Differences, 48*, 517-521.

Whitaker, S. & Wood, C. (2008). The distribution of scale score and possible floor effects on the WISC-III and WAIS-III. *Journal of Applied Research in*

*Intellectual Disabilities*. Available on the Journal of Applied Research in  
Intellectual Disabilities 21, 136-144.