



# University of HUDDERSFIELD

## University of Huddersfield Repository

Whitaker, Simon

Guidance on the Assessment and Diagnosis of Intellectual Disabilities in Adulthood: Getting there but still some way to go

### Original Citation

Whitaker, Simon (2015) Guidance on the Assessment and Diagnosis of Intellectual Disabilities in Adulthood: Getting there but still some way to go. *The Bulletin of the Faculty for People with Intellectual Disabilities*, 13 (3). pp. 16-21. ISSN 2056-3094

This version is available at <http://eprints.hud.ac.uk/26573/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: [E.mailbox@hud.ac.uk](mailto:E.mailbox@hud.ac.uk).

<http://eprints.hud.ac.uk/>

Guidance on the Assessment and Diagnosis of Intellectual Disabilities in  
Adulthood: *Getting there but still some way to go*

By

Simon Whitaker

Consultant Clinical Psychologist/Senior Research Fellow  
Room HWG/16  
University of Huddersfield  
Queen's Gate Huddersfield  
HD1 3DH  
UK  
[s.whitaker@hud.ac.uk](mailto:s.whitaker@hud.ac.uk)

in *The Bulletin of the Faculty for People with Intellectual Disabilities*, 13,  
(3), 16-21.

## **Abstract**

The new BPS *Guidance on the Assessment and Diagnosis of Intellectual Disabilities in Adulthood* are reviewed. It is argued that the continuation of IQ cutoff points and the introduction of adaptive behaviour cut off points are not justified if current research on our ability to measure these two variables is properly considered. The greater emphasis on clinical judgment is welcomed.

The new BPS "Guidance on the Assessment and Diagnosis of Intellectual Disabilities in Adulthood", published in 2015 (BPS 2015), produced by a working group of the Faculty for People with Intellectual Disabilities, replaces the old guidance (BPS 2000). The need for an update seems to have been precipitated by of a number of factors:

- The three international recognised diagnostic definitions of Intellectual Disability (ID) (AAIDD-11, DSM-V and ICD-11) have all been recently updated or are about to be updated. One of the stated aims of the working group was to have a UK definition of ID in line with these definitions.
- Recent UK legislation, notably the Mental Health Act 2007, and the Mental Capacity Act 2005, may lead to psychologists being asked to assess individuals suspected of having an intellectual disability.
- There has been debate in the literature as to how a definition should be framed, in particular whether a measured IQ score should be part of the definition.

The working group tried not to work in isolation, consulting with faculty members and holding two conferences in 2013, in London and Glasgow. The views of members were sought and critics of the current system were able to air their views at the conferences. However, although resulting guidance does differ from the old one, for example, the possibility that IQ tests are not as accurate as we previously thought is acknowledged, the collection of background information is emphasised and the use of clinical judgment is given more prominence, this is mainly a change in emphasis rather than substance. It seems to the current author that, although the definition of ID adopted is consistent with internationally recognised definitions, the working group has not taken sufficient account of recent work on how accurately low IQ and adaptive

functioning can be measured and has missed an opportunity to produce radically different and more effective definition and guidance.

### **The definition of ID**

The definition of ID used in the new guidance is essentially the same as in the old ones and is very much in line with the internationally recognised definition outlined in AAIDD-11, DSM-V and ICD-10/11:

"Intellectual disability is defined as significant impairment in intellectual functioning and significant impairment in adaptive functioning (social functioning), with each of these impairments beginning prior to adulthood." (P 2).

Further details as to what is meant by significant impairments in intellectual and adaptive functioning and age of onset are then given.

A significant impairment in intellectual functioning is "best represented by an IQ score derived from an appropriately standardised and norm-referenced assessment measure that is more than two SDs below the population mean, allowing for the expected level of measurement error within the test" (P 12-13).

A significant impairment in adaptive functioning (social functioning), in line with AAIDD-11, is defined as: "best represented by a score derived from an appropriately standardized and norm-referenced assessment measure that encompasses either: (i) any one; or (ii) all three of the domains of conceptual, practical, and social skills that is more than two SDs below the population mean, allowing for the expected level of measurement error within the test concerned" (P. 14).

With regard to the requirement the disabilities should appear before adulthood, it says "This disability originates before age 18" (P. 14).

The most obvious problems with this definition are as follows:

### **IQ cut off points are still part of the definition**

Having an IQ below 2 SDs or IQ 70 is still a part of the definition. Although this is consistent with the internationally recognised definition of ID there are a number of issues with it that have recently been highlighted in the literature (e.g. Whitaker 2008a, 2013 Webb and Whitaker 2012).

First, IQ 70 or 2 SDs below the mean, is an arbitrary figure, chosen as it is two nice whole numbers (70 and 2), rather than because there is any empirical evidence that people with measured IQs less than 70 are particularly less able to cope than those with measured IQs between 75 and 80.

Second, there is now evidence that IQ tests, partially when used in the low range, are not as accurate as the four to five points that their manuals claim they are. This means that categorising individuals as having a true intellectual ability either above or below 70, on the basis of a measured IQ, is often going to be wrong (see Whitaker 2015 for more details of this). On the one hand, the working group were aware of the possibility of this with the guidance acknowledging that there is evidence that the WAIS-IV may not be as accurate when used in the low range as in the average range (Whitaker 2012), and drawing a distinction between IQ and "true IQ", implying that they are not always the same. However, on the other hand, the guidance does not give any indication as to the degree of this inaccuracy or take account of it in the definition.

It may be helpful to briefly outline the evidence for the lack of IQ tests in the low range:

- The 95% confidence interval (the interval either side of the measured IQ where the "true IQ" supposedly has a 95% chance of lying), which is recommended in the guidance and the test manuals, is mainly based on the internal consistency of the test and not the stability of the test (the degree to which a test will give the same result if given more than once to the same person). This means that this 95% confidence interval does not take account of the change in scores that occur when the test was repeated. In a meta analysis of test re-test reliability of IQ assessments in the low range, Whitaker (2008b) found that the weighted mean test re test reliability score was .82, equivalent to a 95% confidence interval of 12.5 points for stability, and that 14% of re-tested IQs differed by 10 points or more from the original. It would seem to be disingenuous to state that the true IQ is within four points of the measured IQ (which is what is stated in the WAIS-IV manual) if there is a .14 chance that, if the assessment was repeated the next day, the new score would differ by 10 or more points from the original one.

- There is now evidence that commonly used IQ tests don't agree with each other in the low IQ range. Gordon et al (2010) found the WISC-IV measured on average 12 points less than the WAIS-III with to 16-year-olds in special education. Silverman et al (2010) compared the Stanford Binet assessments and WAIS assessments with adults with ID and found that the WAIS measured higher by nearly 17 points. Grondhuis and Mulick (2013) found that the Leiter-R measured higher than the Stanford Binet-V by an average of 21 points when given to children with autistic spectrum disorders. Although the guidance does acknowledge that there may be some problems with the accuracy of the Leiter, there is no indication that the disparity between IQ tests is as great as the evidence suggests it is. As the guidance fails to make this clear it may be the case that many psychologists are unaware of this and will assume that a score on one IQ test is equivalent to the score on another IQ test. Even if there is a general awareness the situation, at the moment we don't know which assessments are accurate and which are in error. This clearly matters when one is trying to estimate an individual's true level of intellectual ability in order to make a diagnosis. It has been suggested (Whitaker 2015) that one way of reducing this problem, though certainly not eliminating it, would be to make the IQ part of the definition explicitly a WAIS-IV IQ. Then if other tests are used the scores on these tests could be adjusted so they are equivalent to WAIS-IV IQs. However, for this to work effectively we would need a lot more data about how other tests compare with the WAIS-IV in the low range, which we don't have at the moment.

Whitaker (2013) has suggested that, when these and other errors are taken into account, the true confidence interval, in the low range, is more like 20 points rather than the 4 to 5 points claimed in the test manuals: a level of accuracy far too small to have a stated IQ cut off point as part of the definition of ID.

### **Adaptive behaviour cut-off points have been introduced**

It is a requirement of the definition that the individual should score 2SDs below the population mean on at least one of: conceptual, practical, or social skills subscales of adaptive behaviour or on the composite scale. This therefore requires an assessment that has been normed against the general population and has normally distributed scores. The guidance recommends the Vineland-II or the ABAS-II are used, both of which were

normed on a stratified sample of the population and give standard scores with a mean of 100 and SD of 15, and so theoretically can be used to assesses if an individual scores above or below the cut off point at 2SDs below the population mean. However, there are a number of reasons why the use of these scales with a 2 SD cut off point may fail to correctly categorise individuals who are able to cope in the real world from those who are not able to:

First, as with IQ, 2 SDs below the mean is an arbitrary figure. There is no empirical evidence that the current author is aware of that people who score 2 SDs below the norm on a scale of adaptive behaviour are any less likely to cope in the real world than those who score at either 2.5 SDs or 1.5 SDs below the population mean.

Second, adaptive behaviour scales seem to be even less accurate than IQ tests (Whitaker 2013). These scales are subject to the same chance errors as are IQ scales, such as a lack of internal consistency and a lack of stability but there is an additional sources of error inherent in having to use an informant to provide information about the individual being assessed. The degree to which the informant knows the individual will vary, the setting in which they have seen the individual may be limited and they may have personal reasons for wanting to present the individual as either capable or incapable. Whitaker (2013) has estimated the true 95% confidence interval of a composite measure of adaptive behaviour, when all relevant sources of chance error are taken into account, is about 18 points. The tests are, however, not only subject to chance error: there are also a number of systematic errors: a floor effect on the Vineland-II causing it to produce a higher score for children than for adults in the low range; there are some differences between the scores on different measures of adaptive behaviour; there is a systematic difference between the rating of parents and teachers, with parents tending to rate their children as more capable (Harrison 1987) and there is a possibility of people deliberately faking scores which is difficult to detect (Doane and Salekin 2009). Acknowledging that it is difficult to calculate an exact interval, Whitaker (2013) estimates the true 95% confidence interval, for adaptive behaviour scales in the measurement of adaptive behaviour in the low range, when both chance and systematic error is taken into account, is of the order of about 20 above the measured score for both adults and children, 20 points below for adults and 30 points below for children. Although the guidance does note that there may be problems with the personal bias of respondents it does not indicate the degree of error that may occur when it comes to assessing an individual's "true" level of adaptive behaviour. As with IQ this error would seem to be far too great to base a definition of ID on the measured score and a large proportion of individuals who

are allocated as having a true level of adaptive behaviour either above or below 2 SDs below the mean will be wrong.

Third, as the scores on adaptive behaviour scales is an average of the individual items it may not indicate when an individual has a key deficit that prevents them from functioning adequately. Greenspan et al (2011) point out that a key element of ID is naivety, which can result in the individual being taken advantage of by others, for example to get involved in crime, lend money etc. This may not be reflected in an adaptive behaviour score if the individual has good linguistic, social and practical skills, but could stop him/her being able to cope adequately. If the definition of ID in the guidance is applied strictly when it comes to adaptive behaviour there may well be individuals who have both measured IQs and true intellectual abilities below 70, who are vulnerable to exploitation but are nonetheless refused a diagnosis and subsequent ID services as they score above the adaptive behaviour cut-off points.

**The definition requires that an individual should fulfil each of the three criteria**

The guidance is quite clear that for a diagnosis of ID the individual should fulfil each of the three criteria. Although this is consistent with the internationally recognised definitions of ID there are some problems inherent in this:

- The age criteria of 18 is largely arbitrary and there may well be cases where for clinical reasons it makes more sense for it not to be applied strictly. For example, if an adolescent receives neurological damage resulting in them meeting the other two criteria but has a very mixed picture of cognitive abilities, they may well be better seen by neurological rehabilitation services rather than ID services.
- The requirement that an individual should have both a significant impairment intellectual ability and adaptive behaviour will mean that individuals who have a true intellectual ability below 70 but a measured adaptive behaviour level in the low normal range would not be diagnosed as ID. As noted above, this could result in some individuals who need an ID service not being given one if they fail to fulfil the adaptive behaviour part of the definition, even if they have a true intellectual ability below 70. This may be a particular problem with regard to service provision of psychological therapies. There is evidence that CBT needs to be modified in order to be effective with individuals with intellectual abilities below 70



(Whitaker 2001), however, if an individual with a true intellectual ability below 70 fails to get a diagnosis due to scoring too highly on a measure of adaptive behaviour then they very well may not be seen by a psychologist specialising in working with individuals with ID.

### **Other issues with the guidance**

There are a number of other points in the guidance that may cause confusion to a psychologist trying to apply the recommendations.

### **Reasons for inaccurate scores**

The guide states that although there are a lot of factors that may reduce IQ scores the only major factor that can increase scores is teaching the test (coaching) for verbal subtests and too short a test-retest interval for performance (non-verbal) subsets (a practice effect). The implication of this is that, in the absence of evidence for coaching or practice, that, if there are two IQ scores, then the higher one is most likely to be accurate. However, the working group may have failed to consider other factors that may result in a score being artificially increased, which would include:

- The Flynn effect, whereby older IQ tests, measure higher than new ones by the order of 0.3 of a point per year since the test was standardized (Flynn 2007).
- The Floor effect, which would increase scores when an individual's true ability on a subtest is lower than the minimum scaled score of one, and/or when their true IQ is below the minimum IQ of 40 (Whitaker and Wood 2008, Whitaker and Gordon 2012).
- Other factors causing systematic difference between tests in the low range that we know occur. For example, as noted above, the WAIS-III/IV may measure too high by the order of 10 points, for reasons we as yet don't understand.
- A higher score may occur by chance. Simply taking account of the 95% confidence intervals of about 4 points reported in the WAIS-IV test manuals, 2.5% of scores will be above this interval. This, however, does not take into account the lack of stability of the test which if included would take the 95% confidence interval to about 7 points for the WAIS-IV (Whitaker 2013). Therefore one would expect that 2.5% of the standardization sample, who would have taken the test under optimal conditions, to score 7 points or more above their true IQs.

### **Reporting scores**

There seems to be some inconsistency in the advice the guide gives on how scores are to be reported. On page 28 it states: "The Society is mindful of a tendency amongst non-psychologists to misinterpret the results of psychometric assessment measures presented in the form of single, standardised scores (such as IQs). The Society, therefore, strongly recommends that neither individual IQ scores nor their equivalents in respect to adaptive behaviour measures should be cited in psychological assessment reports." This is to be welcomed as it is certainly the experience of the current author that non-psychologists can focus on the stated score and ignore everything else in the report. However the guide then says: "the psychologist's opinion as to the individual's *true* level of both intellectual functioning and adaptive behaviour should be presented and discussed. This will involve detailing not just the actual ranges of scores derived throughout the assessment, but also the extent to which the psychologist, using his/her clinical judgment." (P.29). This suggests that scores of some sort should be presented. There is also a specific issue with citing a range of scores instead of the measured IQ score, in that, as shown above, the 95% confidence intervals provided in the manuals are a gross underestimate of the true 95% confidence interval (Whitaker 2008a; 2013).

### **Clinical judgment**

There is an increased emphasis on clinical judgment, which is to be welcomed. The need to interpret and explain the results of the assessments is important, given the problems with the accuracy of tests and how scores can be interpreted by non-psychologists. It has been suggested (Webb and Whitaker 2012, Whitaker 2008b, 2013) that, due to the inaccuracy of measurement, IQ and adaptive behaviour cut off points should no longer be part of the definition of ID and that the decision as to whether an individual has an ID should be largely based on clinical judgment. The guide does not go that far, cut off points are retained and it is implicit that the guidance considers that the measures of IQ and adaptive behaviour are as accurate as is claimed in their manuals. The use of clinical judgment seems therefore to be confined to situations where scores fall within the 95% confidence interval according to the test manuals and where there are more exceptional circumstances, for example where a formal IQ test cannot be given. However, the wording of the guidance seems sufficiently ambiguous, so that clinicians who are aware of the poor accuracy of measures of IQ and adaptive

behaviour can use clinical judgment over a greater range of measured IQs and adaptive behaviour scores.

### **Conclusions**

To conclude, the guidance does have some things to commend it, with the increased emphasis on adaptive behaviour, and clinical judgment and in acknowledgement that there may be problems with the measurement of low IQ. However, the definition of ID adopted is very much in line with internationally recognized definitions, which are based on assumptions about how accurately an individual's intellectual ability and level of adaptive behaviour can be measured which are increasingly being shown to be wrong. It is a pity that the working group did not take the opportunity to produce a radically different definition of ID based on clinical judgment.

### **References**

Doane, B.M. & Salekin, K.L. (2009). Susceptibility of current adaptive behavior to feigned deficits. *Law and Human Behavior*, 33, 329-343.

Flynn (2007). *What is Intelligence: Beyond the Flynn Effect*. Cambridge: Cambridge University Press.

Gordon, S., Duff, S. Davison, 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 Disability*, 23, 197-200.

Greenspan, S. Switzky, H.N. & Woods, G.W. (2011). Intelligence involves risk-awareness and intellectual disability involves risk-unawareness: Implications of a theory of common sense. *Journal on Intellectual & Developmental Disability*, 36 (4), 246-257.

Grondhuis, S.N. & Mulick, J.A. (2013). Comparison of the Leiter International Performance Scale—Revised and the Stanford-Binet Intelligence Scales, 5th Edition, in Children with Autism Spectrum Disorders. *American Journal on Intellectual and Developmental Disabilities*: January 2013, Vol. 118, No. 1, pp. 44-54.

Harrison, P.L. (1987). Research with adaptive behavior scales. *Journal of Special Education*, 21, 37-68.

Silverman, W., Mizejeski, C., Ryan, R., Zigman, W., Krinsky-McHale, S & Urv, T. (2010). Stanford-Binet and WAIS IQ differences and their implications for adults with intellectual disability (aka mental retardation). *Intelligence*, 38, 242-248.

Webb, J. & Whitaker, S. (2012). Defining learning disability. *The Psychologist*, 25, 440-443.

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

Whitaker, S. (2008b). Intellectual disability: A concept in need of revision. *British Journal of Developmental Disabilities*, 54, 3-9.

Whitaker, S. (2001). Anger control for people with learning disabilities: A critical review. *Behaviour and Cognitive Psychotherapy*, 29, 277-293.

Whitaker, S. (2013). *Intellectual Disability: An Inability to Cope with an Intellectually Demanding World*. Basingstoke: Palgrave MacMillan.

Whitaker, S. (2015). How accurate are modern IQ tests at categorising people as ID or non ID? *Clinical Psychology Forum*, 270, 49-53.

Whitaker, S. and Gordon, S. (2012). Floor effects on the WISC-IV. *International Journal of Developmental Disabilities*, 58, 111-119.

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*. 21, 136-141.