



University of HUDDERSFIELD

University of Huddersfield Repository

Whitaker, Simon and Laird, Carmen

Error in the measurement of low intellectual ability:Implications for research

Original Citation

Whitaker, Simon and Laird, Carmen (2011) Error in the measurement of low intellectual ability:Implications for research. In: Joint Congress of the European Association for Mental Health in Intellectual Disability & IASSID Challenging Behaviour & Mental Health SIRG Conference, 1st - 3rd September 2011, Manchester, UK. (Unpublished)

This version is available at <http://eprints.hud.ac.uk/id/eprint/11423/>

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.

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

Error in the measurement of low
intellectual ability:
Implications for research

By

Dr Simon Whitaker
and
Carmen Laird

School of Human and Health Science
University of Huddersfield

Definition of Learning Disability

- An IQ less than 70.
- A deficit in adaptive skills.
- All occurring before the age of 18 years.

IQ Assessments

Focus on the commonly used IQ tests: the WISC-III/IV and the WAIS-III

These are probably the most well standardised and researched psychological tests ever produced.

All tests are subject to error.

An indication of the degree of accuracy of the WISC-IV and WAIS-III is given in the test manuals.

It is claimed that the obtained IQ will be within 5 points of the true IQ on 95% of assessments.

Sources of error in the measurement of IQ

Chance errors:

- Lack of internal consistency.
- Temporal error.
- Scorer error.

Systematic error:

- Flynn effect.
- Floor effect (low range only).
- Lack of consistency between tests.

Chance Error

95% confidence interval

If the degree of chance error is known then a 95% confidence interval (95% CI) can be calculated by:

$$95\% \text{ CI} = 1.96 \times \text{SD} \times \sqrt{(1-r)}$$

SD is the standard deviation of the test and r is the reliability coefficient.

It is usually reported to be about 5 points

The WISC and WAIS 95% confidence interval

- It is based on the performance of the standardization sample, who on the whole had average IQs so may not be representative of people with low IQs
- It is based on the error due to a lack of internal consistency only and does not take into account other sources of error.

Chance Error in low range

Lack of internal consistency

Wechsler (2008) in the WAIS-IV manual. Given to 75 adults with mild ID and 35 with mod. The internal consistency was about .98 which gives a 95% confidence interval of about 4 points.

Temporal Error

An estimate of this is given by the test re-test reliability check.

The test re-test reliability is the correlation between the IQ scores obtained by a group of people being given the same test on two occasions.

A meta-analysis

A meta-analysis of the literature on the stability of intelligence tests when applied to people with low intellectual ability ($IQ < 80$) was done (Whitaker 2008). The mean correlation between first and second test was 0.82.

This corresponds to a 95% confidence interval of 12.47 points.

It was also found that 14% of IQs change by 10 points or more between the two assessments. Which is close to what a 95% confidence interval of 12.5 would predict.

Combining error

A measure of lack of internal consistency does not include temporal error.

A measure of temporal error does not include internal consistency but may include score error.

Error due to lack of internal consistency in low range is $1 - .98$ (Wechsler 2008) = .02.

Error due to temporal changes is $1 - .82$ (Whitaker 2008) = .18

Total chance error is .20.

A total chance error of .20 gives a 95% confidence interval of 13 points for “true IQ”.

Systematic error

The floor effect

The tests measure IQ by giving a client a number of subtests on which there will be a raw score. This raw score is then converted to a scaled score between 1 and 19. The mean scale score is 10 and SD 3.

A scaled score of 1 is given even if the client gets a raw score of 0. This leads to the possibility of an overestimate of IQ.

Scaled scores of 1 may therefore be overestimates of a client's true ability on that subtest.

- Both WISC-III and WAIS-III will be subject to a Floor Effect for IQs in the 40s and 50s but the WISC-III will also be subject to one for IQs in the 70s.
- 10% of scaled scores on the WISC-III for IQs in the 70s were scaled score 1.

The Flynn Effect

Flynn (1984) found that, in the US, the longer it was since the test was standardised the higher the IQ, the rate of increase being about 3 points a decade. The implication is that the intellectual ability of the population as a whole is increasing at a rate of about 3 points a decade or 0.3 of a point per year.

Therefore on average an IQ test will measure about .3 of a point too high for each year since it was standardised.

This would be no problem if the Flynn Effect occurred at a constant rate but its rate has varied over the years.

Lack of agreement between tests

There is evidence that earlier versions of the WISC and the WAIS did not agree at the low IQ level.

We (Gordon, Duff, Davison and Whitaker in 2010) therefore compared the WISC-IV and the WAIS-III in an empirical study on seventeen 16-year-olds in special education.

Results

	WISC-IV	WAIS-III	dif	r
FS IQ	53.00	64.82	11.82	.93

Implications for scientific literature

- The chance error will reduce correlations and the significance levels of statistical tests.
- There is only a loose relationship between an IQ on one test and that on another.
- Simply stating an IQ score, without saying how it was obtained, is almost meaningless.

How is IQ used in the scientific literature?

We read all the papers in JARID and AJMR for 2008.

- 91 papers in the two journals.
- 81 referred to the concept of intelligence.

- Several papers cited IQ scores without saying how they were obtained.
- Several papers assumed that different tests gave the same result.
- Several papers cited very low IQs without saying how they were obtained.

Russell et al (1997)

Follow-up 34 children who had been seen in clinic and had IQ done (mainly on the WISC-R) as adults, after they had developed schizophrenia.

As children assessed on the
WISC/WISC-R. Mean IQ 84.1.

As adults with Schizophrenia assessed on
the WAIS-R. Mean IQ 82.2.

Conclusion Schizophrenia does not
reduce IQ.