The US death penalty and the implication of measurement of low IQ

By

Dr Simon Whitaker

School of Human and Health Science
University of Huddersfield

Between 1977 and 2006 thirty three states in the USA executed 1003 people, of whom an estimated 44 had mental retardation (Patton and Keyes 2006).

In June 2002, the Supreme Court of the United States, in the case of Atkins v. Virginia, made the decision to ban the execution of people with mental retardation.

The court, however, did not give guidance as to a clear definition of mental retardation and left it up to individual states to produce their own. This has resulted in a range of different definitions of mental retardation (Duvall and Morris 2006; DeMatteo et al 2007).

On the 23 September 2010 Teresa Lewis was executed in Virginia.

She had a reported IQ of 72.

Definition of Learning Disability

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

All the 38 states that allow capital punishment require a sub-average level of intellectual functioning as part of their definition of mental retardation. Twenty of these states specify an IQ figure above which a convicted individual would not be considered to have mental retardation and so could not be reprieved from execution on the grounds of mental retardation.

A key question

How accurately can we measure IQ in the low range?

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:

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

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

Davis (1966) who found split-half reliabilities of .90 for children with moderate ID (mean IQ 48) and .97 for those with borderline mental ID (mean IQ 76), the weighted mean reliability being .92 which gives a 95% confidence interval of 8.3 points.

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 scored 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

- There is a lot of scope for error in assessment of mental retardation/ intellectual disabilities.
- It is posible that people with true IQs less than 70 are still being executed.