The British Journal of Developmental Disabilities
Vol. 56, Part 1, JANUARY 2010, No. 110, pp. 49-55

ARE PEOPLE WITH INTELLECTUAL DISABILITIES GETTING MORE OR LESS INTELLIGENT?

Simon Whitaker

Introduction

The Flynn Effect

There is very good evidence that the intellectual ability of the population as a whole is increasing from one generation to the next. In a now classic paper, Flynn (1984) looked at US studies in which the same people had been given two different IQ tests. He found that there was a clear relationship between the time since the test was standardised and the IQ obtained. The longer it was since the test was standardised the higher the IQ, the rate of increase being about three points a decade, the implications being that as tests go out of date they become easier for the population as a whole. In a second paper, Flynn (1987) extended his analysis to 14 industrialised countries and found evidence of an increase in IQ in all of them. For some countries (e.g. Holland, Belgium and France) the data was very strong, being based on the assessment of virtually all 18-year-old men when they report for military service.

This increase in intellectual ability has also occurred in the low IQ range. In a review of studies in which the Wechsler Intelligence Scale for Children (WISC Wechsler 1949) and Wechsler Intelligence Scale for Children – Revised (WISC-R Wechsler 1974) had been given to the same children, Flynn (1985) found that the gains appeared to be higher at the low levels: 0.396 per year for IQs 55 to 70 compared to .272 per year for IQs in the range 125-140. In a more up-to-date review (Flynn, 2006a) he suggests that low IQs are still increasing by about 0.3 of a point per year in the US. Data from the assessment of military conscripts also suggests that the increase in IQ occurs at the low level of intellectual ability. In Norway, military service is compulsory for every able young man, who, as part of his induction process, is given an IQ test. Sundet, Barlaug and Torjussen (2004) used this data to compare the gains made for

***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: +44 (0)1484 473612 Email: s.whitaker@hud.ac.uk

* *For Correspondence*

conscripts scoring above and below the median for pooled data from 1957 to 1959 with data from 1993 to 2002. For those scoring below the median there was an 11 point IQ point gain, which compared to a 4.4 point gain for those above the median. Teasdale and Owen (1989) used similar data from Denmark and found average gains in IQ over the 30 years up to the late 1980s of about 7.5 IQ points. The gains were greatest in the lowest 10%: the maximum gains were near the 11th percentile, at which point the gains were 41% greater than those at the median. At the 90th percentile there was very little gain over the years. However, Teasdale and Owen (2005) looked at the new data up to 2004 and found that there was a peak in average IQ in 1998 and then a decline until 2004. They also report that after 1995 there was an increased number of people scoring at the lower end of the tests showing a decline in IQ for people with lower IQ. There is, therefore, some evidence that the gains in IQ in the low range have stopped or even gone into reverse.

From a UK point of view, one should be cautious of this data, as it relates mainly to the US and Scandinavia. The Scandinavian data, though very good from the point of view of sample size, only relates to men aged between 18 and 19. It is not known what is happening with regard to changes in intellectual ability in the UK and to a large extent what happens at different age groups. One possible source of up to date data on the UK populations is available from the UK standardisations of Wechsler Intelligence Scale for Children – third edition (WISC- III Wechsler 1992) and the Wechsler Intelligence Scale for Children – fourth edition (WISC-1V; Wechsler 2004).

The WISC-III (UK) was standardised between March and July 1991 using a sample of 814 children: 407 boys and 407 girls from 61 schools in the UK. The WISC-IV (UK) was standardised between November 2003 and January 2004 using a sample of 780: 368 boys (47.2%) and 412 girls (52.8%) from 68 UK schools. Both samples were stratified on race/ ethnic group, and geographical region. The WISC-III was also stratified on socio-economic status of parents and the WISC-IV on the educational level of parents. Both claimed to be representative samples of the UK populations: the WISC-III matching the 1989 census and the WISC-IV the 2001 census. In both cases, children receiving special needs support in the schools were not excluded from testing, however no special schools were involved in the study. The two samples may therefore be considered to be equivalent and any changes in the intellectual ability of the samples can be considered to show changes in the population as a whole. If this is the case then any systematic changes in performance on common parts on these two tests will be due to a genuine change in intellectual ability of children over the 12.5 years between the two assessments being standardised.

Although there are a number of differences between the WISC-III and WISC-IV in terms of the subtests used, the items in some of the subtests and the way the results are reported, two subtests, Coding and Symbol Search, are exactly the same in both tests. In addition, a third subtest, Digit Span, has had a minor change which can be compensated for. The performance of the WISC-III and WISC-IV samples on these three subtests can therefore be used to assess how intellectual ability has changed in the UK.

Method

The analysis

Both the WISC-III and WISC-IV measure IQ and other more specific cognitive abilities by giving the client a number of subtests, each of which measures a different aspect of intellectual ability. The maximum score on each subtest varies from subtest to subtest so that the “raw scores” on different subtests are not equivalent to each other. Raw scores are therefore converted to “scaled scores” which for each subtest has a mean of 10, a standard deviation (SD) of 3 and a range from 1 to 19. The test manuals for both assessments give conversion tables between raw scores and scaled scores on each subtest for 33 four month age groups between the ages of 6 years 0 months and 16 years 11 months.

The scaled scores for each possible raw score were obtained from both the WISC-III and WISC-IV manuals at each age band. The WISC-IV scaled scores were then subtracted from the WISC-III scaled scores. The mean difference between scaled scores was then calculated for each age group. As the actual standardisation was done using samples of children in one year age groups, rather than four month age groups in the tables, the mean differences between scaled scores for each year was calculated. This average change in scaled scores was then multiplied by five to give a score in terms of IQ points for each one year age group between 6 years and 16 years over the 12.5 years between the two assessments being standardized.

Due to the possibility of floor and ceiling effects (c.f. Whitaker, 2005) scaled scores of 1 and 19 were excluded from the analysis. Therefore if either the WISC-III or WISC-IV had a scaled score of 1 or 19 then this difference was not included in the mean differences between scaled scores.

In order to assess the Flynn Effect specifically for children with low intellectual ability and for those with high intellectual ability, the above analysis was repeated using only scaled scores (on the WISC-III) of seven or less and scaled scores (on the WISC-III) of 13 or greater.

The analysis was done on the two subtests, Symbol Search and Coding, which are exactly the same in both the assessments, and Digit Span which is the same on both tests except that on the WISC-IV there is a second two digit item on digits reversed. As, in the author’s experience as a clinical psychologist in intellectual disability, it is very rare for a client, even with a learning disability, not to get the first item in digits reversed correct, it was felt that it could be assumed that everybody in the standardisation sample would have got this item correct and therefore a raw score on the WISC-III was the equivalent of that score plus two on the WISC-IV.

Results

FIGURE 1 shows the mean differences between scaled scores for each year for Symbol Search.

The mean difference between scaled scores overall was equivalent to a gain of 3.43 IQ points. It is notable that the gains were greater at age 6 and 7 years. Those with high IQs had a mean difference equivalent to a gain of 10.43 IQ points. On the other hand, those with low IQs showed a mean negative effect equivalent to a loss of 1.77 IQ points, suggesting that people with low IQs have become less able on this subtest.

FIGURE 2 shows the mean differences between scaled sores for each year for Coding.

There was an overall gain equivalent to 2.39 IQ points over the 12.5 years between the standardisation of both tests. The increase in ability is noticeably greater for those with high IQs, with a mean difference of 6.10 IQ points. With those with low IQs, there was a loss in ability equivalent to 1.46 IQ points.

FIGURE 3 shows the mean differences between scaled scores for each year for Digit Span.

This suggests there has been very little change overall with a mean difference between equivalent to -0.19 IQ points There is a slight increase in ability for those with high IQs equivalent to 0.82 IQ points, and a slight drop for those low IQs with a mean difference equivalent to 0.52 IQ points.

Discussion

For two of the subtests, Symbol Search and Coding, there was a clear increase in ability overall; however, for Digit Span there was very little change. On each subtest there was a greater increase in ability for those with high IQs and a decrease in ability for those with low IQs.

This analysis therefore suggests that the Flynn Effect is continuing for children as a whole; however, it is much greater for children with high IQs and may be going into reverse for those with low IQs. One should, however, be cautious in drawing firm conclusions from this study as there are a number of shortcomings that need to be considered.

The samples on which the tests were standardised were relatively small, only having 74 subjects at each one year age group in the case of the WISC-III, and

Figure 1
The changes in the ability on Symbol Search, in IQ points, between the WISC-III and WISC-IV (WISC-III minus WISC-IV) for all the children in the standardization sample, those with scaled scores less than 8 (on the WISC-III), and those with scaled score of greater than 12 (on the WISC-III), for each year age group between 6 and 16 years.

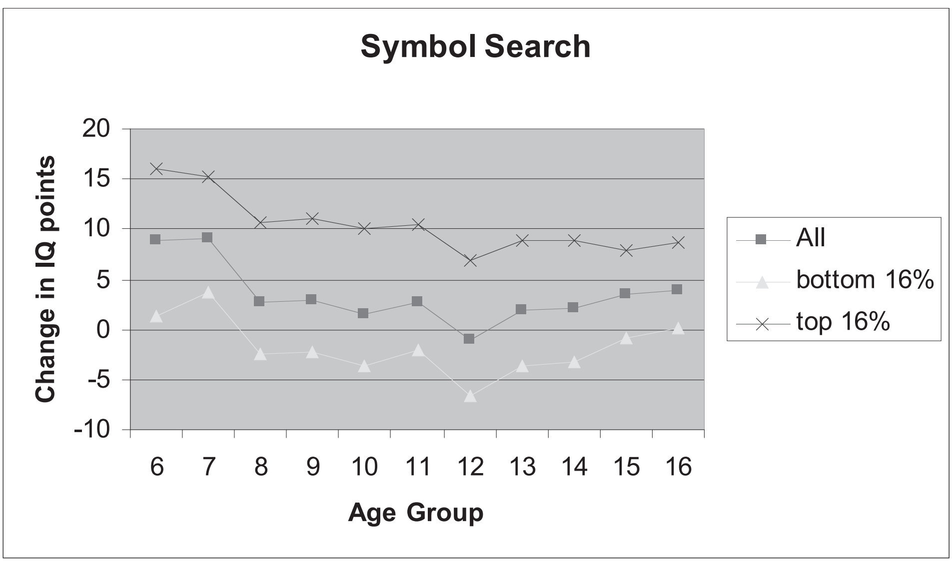


Figure 2
The changes in the ability on Coding, in IQ points, between the WISC-III and WISC-IV (WISC-III minus WISC-IV) for all the children in the standardization sample, those with scaled scores less than 8 (on the WISC-III), and those with scaled score of greater than 12 (on the WISC-III), for each year age group between 6 and 16 years.

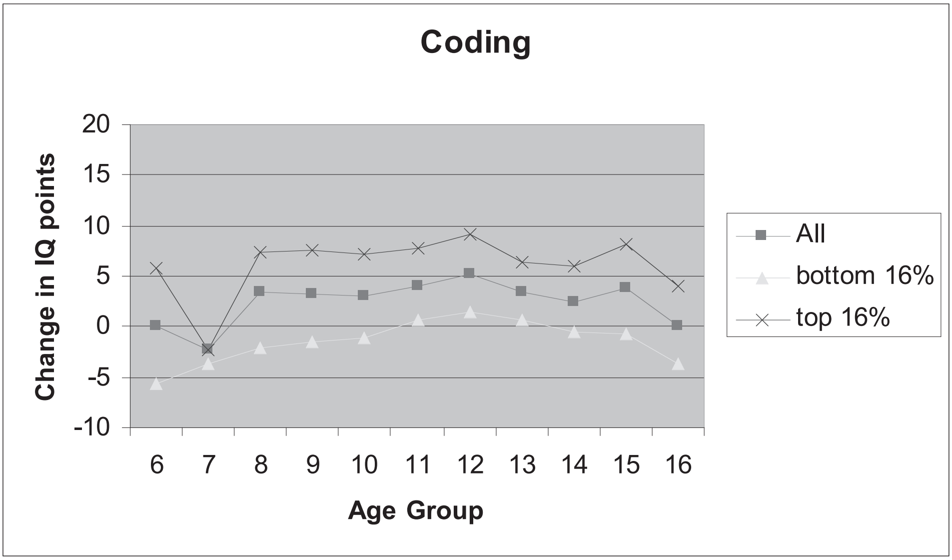
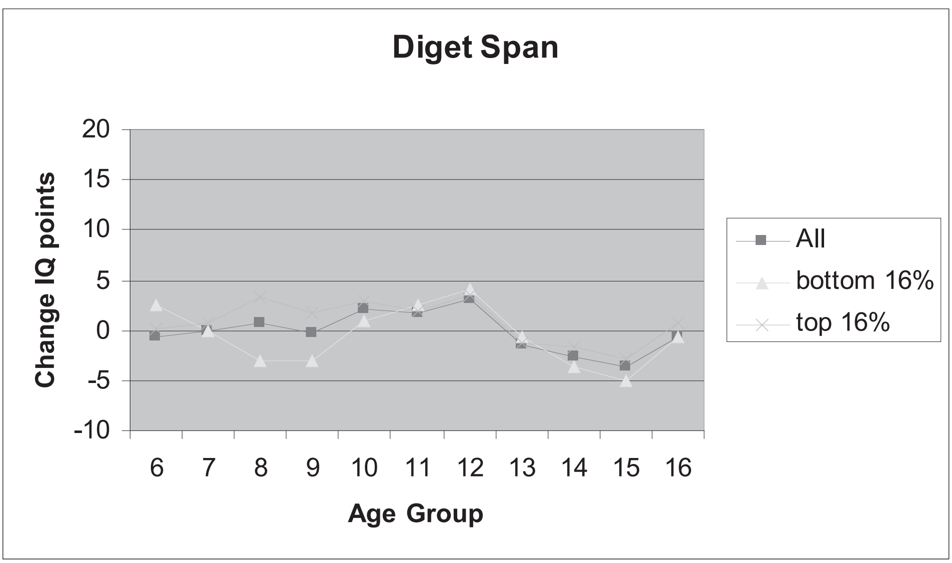


Figure 3
The changes in the ability on Digit Span, in IQ points, between the WISC-III and WISC-IV (WISC-III minus WISC-IV) for all the children in the standardization sample, those with scaled scores less than 8 (on the WISC-III), and those with scaled score of greater than 12 (on the WISC-III), for each year age group between 6 and 16 years.



71 in the case of the WISC-IV. Although it was suggested that the samples were equivalent, this assertion was based on the description of the sampling procedure, which may well have been subject to some error, particularly at the low and high ability levels. Also the analysis is based on only three subtests that were only moderately correlated with Full Scale IQ and so may not be representative of IQ overall. Over the years, the Flynn Effect has been very different for different subtests. Flynn (2006a) notes that the gains have been greatest on assessments of fluid intelligence, notably Similarities and Block Design, and virtually nil on other tests such as Vocabulary and Information. The subtests looked at here are largely tests of fluid intelligence and so may have shown an increase for this reason, while other subtests such as Vocabulary, Information and Comprehension, may have gone into decline.

However, if the results are indicative of what is currently happening to intellectual abilities of children in the UK, they would seem to have implications for people with intellectual disabilities. Although there is no generally agreed mechanism for why the Flynn Effect occurs (see Neisser 1998), Dickens and Flynn (2001) and Flynn (2006b) have proposed an elaborate theory based on a positive feedback loop. They suggest that the more intellectually demanding environment that we now live in requires people to exercise more on-the-spot problem solving, resulting in increased IQ, which in turn results in the environment becoming more cognitively demanding. Clearly the environment has become more intellectually demanding both in terms of work, leisure and caring for oneself. This, according to Flynn (2006b), has resulted in the recent overall gain in intellectual ability. However, the intellectual demands of the environment are determined largely by the intellectually most able members of society. People of average intellectual ability may well be able to learn to cope with these demands and so their IQs are increased. However, it is possible that people with low IQs are not able to adapt to these demands, recognise that they cannot, and so start to avoid intellectually demanding tasks and so become less able intellectually.

Summary

It has been well documented that over the last 60 years there has been a gradual increase in the intellectual ability of the population as a whole which has included people with low intellectual ability. The present study examined if this trend was still continuing in the UK by comparing the scaled scores given for individual raw scores on three common subtests (Symbol Search, Coding, and Digit Span) of the UK versions of the WISC-III and WISC-IV. It was found that over the 12.5 years between the two assessments being standardised there was an overall increase in intellectual ability. However, this increase was greater for those with high intellectual ability. For those with low intellectual ability there was a decline in their intellectual ability.

References

Dickens, W.T. and Flynn, J.R. (2001). Heritability estimates versus large environmental effect: The IQ paradox resolved. *Psychology Review*, 108, 346-369.

Flynn, J.R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29-51.

Flynn, J.R. (1985). Wechsler intelligence tests: Do we really have a criterion of mental retardation? *American Journal of Mental Deficiency*, 90, 236-244.

Flynn, J.R. (1987). Massive gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171-191.

Flynn, J.R. (2006a). Tethering the elephant capital cases, IQ and the Flynn Effect. *Psychology, Public Policy and Law*, 12, 170-189.

Flynn, J.R. (2006b). Efeito Flynn: Repensoado a inteligencia e seus efecitos [The Flynn Effect: Rethinking intelligence and what affects it]. In C Florres-Mendoza & R Colon (eds.) *Introducao a Psicologia das Diferencas Individuais [Introduction to the psychology of individual differences]* (pp. 387-411). Porto Alegre: ArtMed. (English translation from jim.flynn@stonebow.otago.ac.nz)

Neisser, U. (1998). *The Rising Curve: Long-term Gains in IQ and Related Measures*. Washington DC: American Psychological Association.

Sundet, J.M., Barlaug, D.G. and Torjussen, T.M. (2004). The end of the Flynn Effect? A study of secular trends in the mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence*, 32, 249-262.

Teasdale, T.W. and Owen, D.R. (1989). Continuing secular increases in intelligence and stable prevalence of high intelligence levels. *Intelligence*, 13, 255-262.

Teasdale, T.W. and Owen, D.R. (2005). A long-term rise and recent decline in intellectual test performance: the Flynn Effect in reverse. *Personality and Individual Differences*, 39, 837-843.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York: The Psychological Corporation

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children-Revised*. New York: The Psychological Corporation

Wechsler, D. (1992). *Wechsler Intelligence Scale for Children – Third Edition UK*. London: The Psychological Corporation

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children – Fourth UK Edition: Administration and Scoring Manual*. London: The Psychological Corporation

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.