Whitaker, Simon

Are people with Intellectual disabilities getting more or less intelligent II: US data

Original Citation

Whitaker, Simon (2011) Are people with Intellectual disabilities getting more or less intelligent II: US data. British journal of developmental disabilities, 57 (113). pp. 159-166. ISSN 0969-7950

Are people with Intellectual disabilities getting more or less intelligent II: US data

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 347600

s.whitaker@hud.ac.uk

Introduction

The Flynn Effect

There is now good evidence that the intellectual ability of the population as a whole is increasing from one generation to the next (Flynn 1984, 1987, 2006), a phenomenon now known as the Flynn effect after James Flynn, who conducted much of the work on the effect. This increase in intellectual ability has clearly occurred at the low range of intellectual ability, at least in western industrialised countries (Flynn, 1985, 2006, 2009), over at least the last 50 years. However, there is evidence that in Scandinavia it may have stopped or even gone into reverse. In Denmark, military service is compulsory for about 10% to 15% of able young men, chosen at random, who, as part of their induction process, are given an IQ test. Teasdale and Owen (1989) used this data to look for changes in intellectual ability of 18 year old men in 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.

In an attempt to investigate whether intellectual ability of children in the UK was still increasing, Whitaker (2010) used data from the last two UK editions of the Wechsler Intelligence Scale for Children third Edition (WISC-III Wechsler 1992) and the Wechsler Intelligence Scale for Children fourth Edition (WISC-IV Wechsler 2004), which were standardised 12.5 years apart. He compared the scaled scores obtained of each possible raw score on the Symbol Search and Coding subtests, which were exactly the same on both versions of the test. On the Digit Span subtest however, there is a minor difference between the two versions which could be compensated for. He found on the Symbol Search and Coding subtests that, over the full ability range, there was an increase in ability, and that this increase was greater for those in the upper 16% of the intellectual range, but the ability of those in the bottom 16% decreased, suggesting that for those with low intellectual ability the Flynn effect may have gone into reverse. With Digit Span there was effectively no change at any of the ability levels. The results for the full ability range on all three subtests are consistent with what has been reported to for these subtests in the US using the US standardisation of the tests (Flynn and Weiss 2007). However, there have been no reports of a decrease in ability at the low range on Symbol Search or Coding in the US. This possible decrease in intellectual ability in the low range may therefore be confined to the UK. However, there is also the possibility that the results are an artefact of poor standardisation of the WISC-III and WISC-IV in the UK, where the standardisation sample was much smaller than in the US. The WISC-III (UK) was standardised on 814 children and the WISC-IV (UK) was standardised on 780 children. Both the US edition of the WISC-III and WISC-IV were standardised on 2,200 children. Repeating the analysis on the US data would therefore not only give information as to whether the same effect was occurring in the US but also, if it was, give support to the UK findings. The aim of this study was therefore to repeat this analysis on the US data.

Method

The US version of the WISC-III was standardised in 1989, the US WISC-IV in 2002. Both samples were obtained in a similar way and were reported as being representative of US children between 6 and 16 years old. It is therefore a reasonable assumption that the samples were representative of US children when the standardisation was done and therefore any differences between the intellectual performances of the two samples is indicative of a change in the intellectual ability of the US children as a whole between 1989 and 2002.

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 administrative manuals of the tests 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 the US WISC-III and WISC-IV administrative manuals (Wechsler 1991, 2003a), 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 were 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 standardised.

Due to the possibility of floor and ceiling effects (c.f. Whitaker, 2005, Whitaker and Wood, 2008), 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 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, which corresponds to approximately the bottom and top 16% of the samples.

The analysis was done on the three 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 change in ability on the Symbol Search subtest, for each age group between 6 and 16. Change in ability is given on a scale with a standard deviation of 15, equivalent to IQ scores, rather than three, which would be equivalent to scaled scores. This is done to ease interpretation rather than imply that IQ as such is necessarily changing. In order that a comparison can easily be made between the effects in the UK the data from the UK analysis (Whitaker 2010) is also presented. Based on the data from the US manual, the performance of the children in the WISC-IV sample, averaged across all the age bands, showed a 3.96 IQ equivalent point increase on Symbol Search over the WISC-III sample. The effect is greater for the top 16% who showed an increase equivalent to 11.20 IQ points; however, for those in the bottom 16%, there was effectively a change in ability with an average decrease of .02 IQ points. There was not a systematic trend across the age bands with the effect being evident for the full sample of 6 year olds and for those at the high ability level. Comparing the US analysis with that done on the UK shows the effects are similar. The UK analysis on Symbol Search sample overall showed an overall increase equivalent to 3.43 IQ points, an increase of 10.43 IQ equivalent points but a decrease of 1.77 points for the bottom 16% again with no trend across age bands.

Figure 2 shows the equivalent analysis for Coding. The US data showed an overall increase of 1.77 IQ equivalent points of the full sample on average across the age bands, an average increase of 5.50 IQ equivalent points for the top 16% but a decrease of 2.18 IQ equivalent points for the bottom 16%. There may be an upward trend in the data across the age bands from age 8 to 16 for the full sample and for the top 16% though not for the bottom 16%. Again the effect shown by the US data is similar to that found in the UK where there was an increase equivalent to 2.39 IQ points for the sample overall, an increase of 6.10 IQ equivalent points for the top 16% and a decrease of 1.46 equivalent IQ points for the bottom 16%. Although there was no trend across age bands in the UK data, both the US and UK graphs do show an apparent drop in the increased ability for the 7 year age group. It seems unlikely that this drop reflects a sudden change in ability of 7 year olds in both the US and UK on Coding and may well be an artefact of the scoring or standardisation of the tests.

Figure 3 shows the analysis for Digit Span. The US data showed very little difference in ability between the WISC-III and WISC-IV samples for any ability levels or age group. The overall change was equivalent to -.06 IQ points; the top 16% showing an increase of .96 IQ equivalent points and the bottom 16% a decrease equivalent to 1.05 IQ points. However, for the age groups 6, 7, 8 and 9 years, there may have been an effect with the top 16% appearing to show an increase in ability and the bottom 16% a decrease. Again the US data were similar to the UK data where there was a decrease of .19 IQ equivalent points for the full sample across the age bands, an increase of .82 equivalent IQ points for the top 16% and a decrease of .52 equivalent IQ points for the bottom 16%.

Discussion

In brief, the results show an increased ability on Symbol Search and Coding for the full sample, a greater increase for the most intellectually able 16% and either a decrease or no change for the bottom 16%, in both the US and UK. There is no major trend across age groups with the effect being seen in the young children as well as the older ones. For Coding there is effectively no change, except possibly for 6, 7 and 8-year-olds in the US where the pattern of change was similar to the other two subtests. As these results are based on standardisations of the WISC-III and WISC-IV done independently in the US and UK, with the US standardisation having a large sample of 2,200 children they would seem to be robust.

If it is assumed that the effects on Symbol Search and Coding are representative of the Flynn effect in general then it may be possible to draw the following conclusions:

- That the Flynn effect is still occurring for most people in the US and UK,
- it is greater for people with above average intellectual ability
- but has either stopped or gone into reverse for those with lower intellectual ability.

However, it may well be premature to draw these conclusions for a number of reasons:

First, Symbol Search and Coding may not be strong predictors of overall intellectual ability, with correlations with Full Scale IQ (FS IQ) of .57 and .66 for Coding and Symbol Search respectively reported for on the WISC-IV (Wechsler 2003b). Also Symbol Search and Coding are the two core subsets on the WISC-IV that measure Processing Speed Index (PSI) which has the lowest correlates with FS IQ (r=.70) of all four index scores.

Secondly, as they are the only two of the 10 core subtests on the WISC-IV to measure the PSI, it is possible that they have non-intellectual skills in common that are not shared by other subtests. For example, they are the only subtests that require the client to respond using a pen or pencil, so being able to write fast could be a none-intellectual ability that varies, that accounts for the results.

Thirdly, gains on these two subtests will probably be much greater than the gains on most other subtests. Flynn and Weiss (2007), using data from the Psychological Corporation from comparisons of the US versions of the WISC-III and WISC-IV, with subjects in the average intellectual range, found that combined Symbol Search and Coding had a strong Flynn effect of 4.78 equivalent IQ points, significantly above the mean rate of 3.05 for all the subtests. It is therefore clear that the Flynn effect on Symbol Search and Coding may well not be typical of what is happening for FS IQ, either for those of average intellectual ability or those with either high or low intellectual ability. However, it is also clear that, whether the results are due to changes in intellectual ability or non intellectual ability, the change is very different in the upper and lower intellectual ranges and it is legitimate to speculate as to why this may be the case.

A further finding of this study is that the Flynn effect is apparent for 6-year-old children on both Symbol Search and Coding. We are therefore looking for causes that would differently affect those with high intellectual ability compared with those with low intellectual ability, and be present by at least the age of six years. The issue of age at which a Flynn effect is apparent has been addressed by Lynn (2009). He reviewed the evidence for the Flynn effect in babies as young as six months old and infants in the first two years of life and found evidence for it. Based on studies in which different standardisations of developmental scales had been given to the same infants, he found clear evidence that the effect was present for babies as young as six months old. He therefore argued that the effect could not be caused by factors that would only have an impact after infancy, such as improved education and the use of complex leisure activities. He suggests that the most likely reason for the effect was improved diet. Clearly the diet of both child and/or of his/her mother pre-natally may have an influence on a child's intellectual ability, however, if the results of the current study are correct in showing the effect is different at the high and low ability range, it would need to be demonstrated that diet has got better for people with higher intellectual abilities and worse for those with lower intellectual abilities. Another possible factor is family size, suggested by Sundet et al (2008), which is negatively related to IQ and may well have decreased for those with higher IQs and increased for those with lower IQs. A possible mechanism whereby family size has its influence on a child's intellectual ability is that the child will receive more individual parental attention. If some of the additional parental attention received by children in small families involves using paper and pencils and/or playing games that involve fast thinking, such as early computer games, this may be one explanation as to why the Coding and Symbol Search are differentially affected as compared with Digit Span. However, this would not explain why the Flynn effect as a whole was present in infants as young as six months of age as shown by Lynn (2009).

This study raised a number of other research questions: First, whether the differential Flynn effect across ability levels found in this study for Coding and Symbol Search occur in other measures of intellectual ability. A study that could be done would be to give the WISC-III and WISC-IV to separate groups of children with low, average and high intellectual abilities in order to see if the differential effect occurs for the full tests. But the results would have to be interpreted with caution as the WISC-III and WISC-IV are very different tests and the new subtest in the WISC-IV may not be as affected by the Flynn effect to the same extent as the old ones, therefore, meaning a change in FS-IQ would not be a precise measure of the effect. Flynn and Weiss (2007) suggest that comparisons should be made at the level of individual subtests in addition to FS-IQ.

Second, work needs to be done to shed more light on the causes of the effect. Lynn (2009) suggests that whatever is causing intellectual ability to change across the generations is present by six months and the current study suggests that the effect may now be differential for those with high as opposed to low intellectual ability. Possible factors are diet, as suggested by Lynn (2009), family size suggested by Sundet et al (2008), or other environmental factors such as pre-natal care or air or water pollution. Although studies to find answers to this question would not be easy to do their results could have a major influence on the lives of people with intellectual disability as they could point to interventions that could be introduced pre-natally or soon after birth, that could have the affect of boosting a child's intellectual ability for life.

Summary

It has been documented that over the last 60 years there has been a gradual increase in the intellectual ability of the population as a whole including people with low intellectual ability. Using data from the UK standardisations of the WISC-III and WISC-IV, Whitaker (2010) found that for two subtests, Symbol Search and Coding, there was evidence of an increase in ability over the 12.5 years between the standardizations of the two tests. The increase was greater for those with high ability but for those with low intellectual ability there was a reduced ability. A further subtest, Digit Span, showed no change. The current study repeats this analysis for the US standardisations of the WISC-III and WISC-IV and finds essentially the same results.

Acknowledgments

I would like to thank Professor James Flynn for suggesting this study be done and for providing helpful suggestions on the interpretation of the results.

References

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

Flynn, J.R. (2009). The WAIS-III and WAIS-IV: Daubert motions favor the certainly false over the approximately true. Applied Neurology, 16, 98-104

Flynn, R.J. and Weiss, L.G. (2007). American IQ gains from 1932 to 2002: The WISC subtests and educational progress. *International Journal of Testing*, 7, 209-224

Lynn, R. (2009). What has caused the Flynn effect? Secular increase in the developmental Quotient of infants. *Intelligence*, 37, 16-24

Sundet, J.M., Borren, I. and Tambs, K (2008).The Flynn effect partly caused by changing fertility patterns. Intelligence, 36, 183-191

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. (1991). *Wechsler Intelligence Scale for Children – Third Edition*. San Antonio TX: The Psychological Corporation

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

Wechsler, D. (2003a). *Wechsler Intelligence Scale for Children – Fourth Edition: Administration and Scoring Manual*. San Antonio TX: The Psychological Corporation

Wechsler, D. (2003b). *Wechsler Intelligence Scale for Children – Fourth Edition: Technical and Interpretive Manual*. San Antonio TX: The Psychological Corporation.

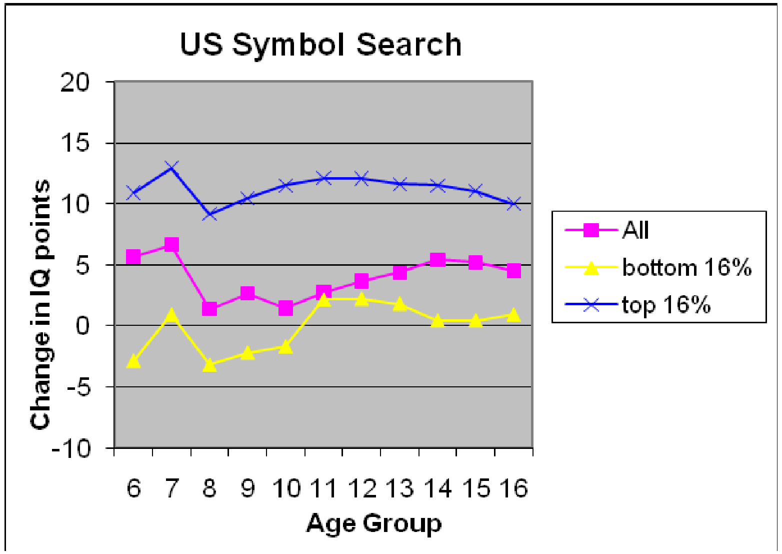
Wechsler, D. (2004). *Wechsler Intelligence Scale for Children – Fourth UK Edition: Administration and Scoring Manual*. London: The Psychological Corporation

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

Whitaker, S. (2010). Are people with intellectual disabilities getting more of less intelligent? The British Journal of Developmental Disabilities, 56, 49-55

Whitaker, S. and Wood, C. (2008). The distribution of scaled scores and possible floor effect on the WISC-III and WAIS-III. *Journal of Applied Research in Intellectual Disabilities*, 21, 136-141.

Figure 1


US Symbol Search
20
15
10
5
0
-5
-10
Change in IQ points
6 7 8 9 10 11 12 13 14 15 16
Age Group
All
bottom 16%
top 16%

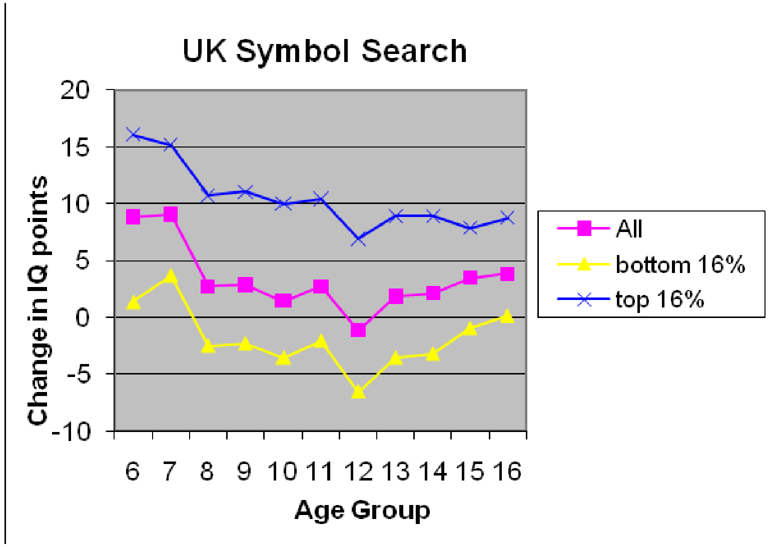
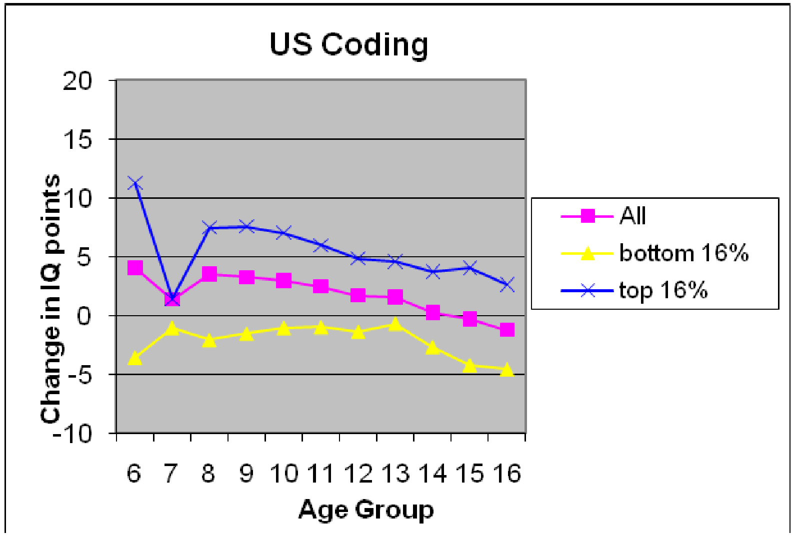

UK Symbol Search
20
15
10
5
0
-5
-10
Change in IQ points
6 7 8 9 10 11 12 13 14 15 16
Age Group
All
bottom 16%
top 16%

Figure 2


US Coding
Change in IQ points
Age Group
All
bottom 16%
top 16%
20
15
10
5
0
-5
-10
6 7 8 9 10 11 12 13 14 15 16

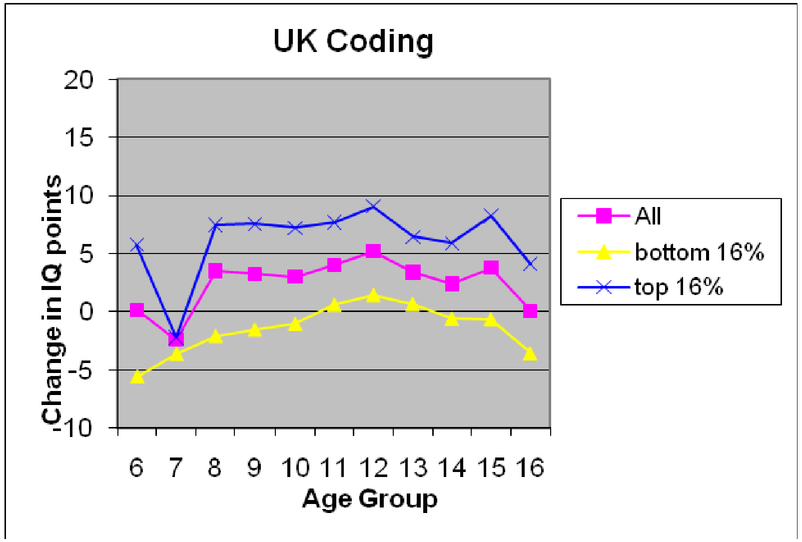

UK Coding
Change in IQ points
Age Group
All
bottom 16%
top 16%
20
15
10
5
0
-5
-10
6 7 8 9 10 11 12 13 14 15 16

Figure 3

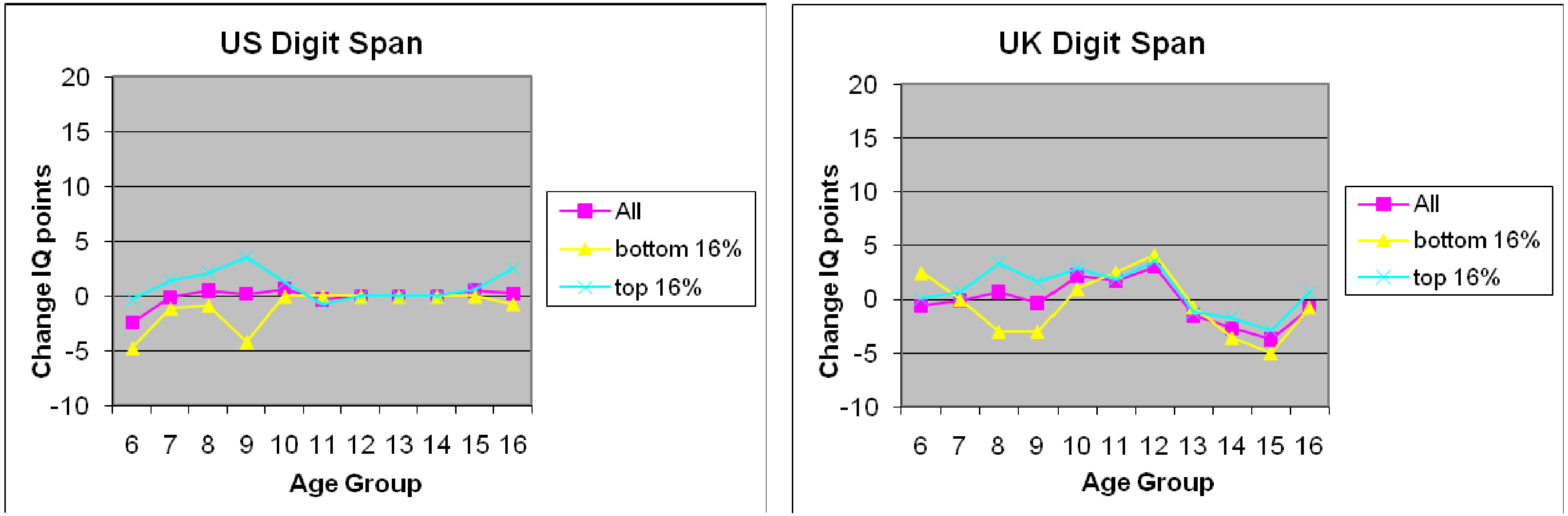

US Digit Span
Change IQ points
Age Group
All
bottom 16%
top 16%
20
15
10
5
0
-5
-10
6 7 8 9 10 11 12 13 14 15 16
UK Digit Span
Change IQ points
Age Group
All
bottom 16%
top 16%
20
15
10
5
0
-5
-10
6 7 8 9 10 11 12 13 14 15 16

Figure captions

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, the bottom 16% (on the WISC-III), and the top 16% (on the WISC-III), for each year age group between 6 and 16 years, for both the US and UK samples.

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, the bottom 16% (on the WISC-III), and the top 16% (on the WISC-III), for each year age group between 6 and 16 years, for both the US and UK samples.

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, the bottom 16% (on the WISC-III), and the top 16% (on the WISC-III), for each year age group between 6 and 16 years, for both the US and UK samples.