Black Americans Reduce the Racial IQ Gap: Evidence from Standardization Samples

Forthcoming: *Psychological Science,* **October 2006**

By William T. Dickens[1] & James R. Flynn[2]

[1] *Economic Studies, Brookings Institution, 1775 Massachusetts Ave. N.W., Washington, DC 20036.*

2 *Department of Political Studies, University of Otago, Dunedin, New Zealand.*

Correspondence: jim.flynn@stonebow.otago.ac.nz

Abstract

It is often asserted that blacks have made no IQ gains on whites, despite relative

environmental gains, and that this adds credibility to the case that the black/white IQ gap

has genetic origins. Until recently, there have been no adequate data to measure black IQ

trends. We analyze data from nine standardization samples for four major tests of

cognitive ability. These suggest that blacks have gained 5 or 6 IQ points on non-Hispanic

whites between 1972 and 2002.   Gains have been fairly uniform across the entire range

of black cognitive ability.

Philippe Rushton and Arthur Jensen (2005) state that the IQ difference between black and

white Americans stands at 1.1 standard deviations (SDs) and is as large today as it was

nearly 100 years ago. No one really knows the history of the black/white gap. Estimates

for 1917 and 1943 are based on military data subject to a host of biases. Estimates since

1945 are based almost entirely on averaging studies, none of which compare nationally

representative samples taking the same test administered at two different times

(Herrnstein & Murray, 1994). Flynn (1987) analyzed military data and found that blacks

gained 3 points on whites between 1940 and 1960 but the estimate is tentative. The racial

IQ gap may have been about 1.1 SDs in the late 1960s. The average from two nationally

representative samples, the 1965 Coleman Report (Jensen, 1980, p. 479) and the 1972

WISC-R (Harcourt, 2005a), is 1.108 SDs.

Data

The inclusion of blacks in recent standardization samples offers better data. We have results from standardizations of four tests: 1972, 1989, and 2002 for the WISC (Wechsler Intelligence Scale for Children), called the WISC-R, WISC-III, and WISC-IV respectively (Harcourt, 2005a); 1978 and 1995 for the WAIS (Wechsler Adult Intelligence Scale), called the WAIS-R and WAIS-III (Harcourt, 2005a); 1980 and 1997 for the Armed Forces Qualification Test, called the AFQT (Department of Defense, 2005); 1985 and 2001 for the Stanford-Binet-4 and the SB-5 (Thorndike, Hagen, & Sattler, 1986, pp. 34-36; Riverside, 2005). Race results are not available for previous standardisations. The publishers of the Wechsler and the Stanford-Binet tests provided sample sizes, mean IQs, and SDs by age group for whites and blacks. The U.S. Department of Defence provided individual test scores, along with information on subjects' age, race, ethnicity, and sample weight for the two standardizations of the AFQT. Summary data are in Appendix A.

The AFQT is not individually administered but is one of the most highly g-loaded tests in use (g refers to the general intelligence factor). It correlates with the classic IQ tests more highly than they do with one another (Herrnstein and Murray, 1994, pp. 580-585). The g measured by the AFQT is skewed toward crystallized g (the tools and skills that intelligent people tend to acquire), but the same is true of the Wechsler tests (Jensen, 1987, p. 96).

Our values for black IQ use the convention that sets the white mean at 100 and SD at 15.   For example, if pre-rise blacks were 1.1 white SDs below whites, this becomes an IQ of 83.5.

Race and samples

Whenever we refer to whites, we mean non-Hispanic whites. Hispanics score below other whites and in recent years, their numbers have dramatically increased. Were they not excluded, the black/white IQ gap would lessen irrespective of black gains.

Two samples do not distinguish Hispanic and non-Hispanic whites: the 1972 WISC-R and the 1978 WAIS-R. We adjusted by raising the white means by 0.70 IQ points and 0.62 points respectively. Appendix A describes how we arrived at these values. The adjustments made little difference. Estimated rates of gain are increased by less than 0.03 IQ points per year, and estimated black IQ at age 12 in 2002 by .022 points.

The number of black/non-black marriages increased from about 1% in 1970 to 4.5% in 1990 (Farley, 1995; Staples, 1985). Thus, an increasing percentage of black children tested between 1980 and 2000 had half-white ancestry.   For the moment, take it on faith that the relevant black/white IQ gap is less than 15 points. If such children scored half way between the black and white means, they would be 7.5 IQ points above the black mean. Arithmetic shows that the increase of such children (up 3.5%) would cause a rise of 0.263 IQ points (.035 x 7.5 = 0.263). Perhaps blacks that can pass for white are more likely to wish to do so today. Even if this is so, the numbers must be very small. A scenario that would maximize the decline in the IQ differential: a child whose

father or mother passed for white decides to declare herself black. If the socially white parents have provided an environment in which their children matched the white mean of 100, each such child would confer a bonus of 15 IQ points. If 1% more blacks fall into that category today, they would cause a rise of 0.15 IQ points.   Shifting group membership is probably not very important.

The black/white IQ gap could be affected by changes in test content. The Wechsler and Stanford-Binet organizations assured us that no item or subtest has been added or deleted to influence the racial IQ gap. Between 1980 and 1997, the AFQT changed from a pencil and paper test to a computerized test. However, in 1997, a large sample was randomly allocated between the two tests. Segall (1997, pp. 192-193) found that the computerized test gave neither black nor white any statistically significant advantage.   Finally, Jensen (1992) has shown that black/white differences tend to be larger on tests that correlate more highly with *g* (the general intelligence factor). The correlation between test scores and *g* rose by 12 per cent from the SB-4 to SB-5 (Roid, 2003, p. 108); rose by 1 per cent from the WISC-R to WISC-IV; declined by 5 percent from the WAIS-R to the WAIS-III (Harcourt, 2005b); and remained the same on the AFQT (Department of Defence, 2005). There is no reason to believe that altered tests became easier for blacks.

The Wechsler and Stanford-Binet manuals show meticulous sampling of schools and weighting to ensure that standardization samples matched census data. Up to age 15, virtually all American children are in school and can be sampled and counted. One qualification: unlike the SB-4, the SB-5 sample included special education and limited English proficiency groups. Because a higher percentage of blacks than whites are in

these categories, this would lower the SB-5 black mean and deflate the SB estimate of black IQ gains.

Adults pose sampling problems but individual data on the AFQT allows a test of their significance. Neal (in press) finds that the 1980 National Longitudinal Survey of Youth (NLSY) sample contained a sizeable group of under-educated blacks that could not cope with the AFQT.   No such group was present in the 1997 NLSY sample, so a comparison of the two would over estimate black gains. Our results are based not on NLYS samples but on the Profile of American Youth (PAY) samples used to norm the AFQT. However, it is desirable to ensure that no such bias is present in these.   If it were, we should see disproportionately large gains among the lowest scoring blacks. In addition, over the years, more and more blacks may have become too isolated to locate. Therefore, fewer below-average blacks might be present in the 1997 (PAY) sample and the census data against which the sample was weighted.   This would mean higher gains below the median tailing off above.

Insert Figure 1 here

Figure 1 presents black IQ gains on whites at each percentile of the black AFQT distribution. Gains are relatively uniform across the entire distribution of black ability, at least from the 3rd to the 88th percentiles. Only the bottom 2% showing heightened gains and even here, the difference is small. Another check: more blacks (mainly young males) have been incarcerated between 1980 and 1997 and might have escaped the later sample. We reviewed the NLSY data for a marked rise in the back/white gap at the ages of heavy incarceration and found nothing.

Trends from standardization samples

Figure 2 shows that black IQ rose on each of our four tests. Appendix B gives the procedure used to construct average black IQs. All but the WAIS cover ages under 25 (the WISC 6 to 16, the SB 3 to 23 and the AFQT 18 to 23). The WAIS solid line traces its trend for those under 25 and the WAIS broken line its trend for all ages (16 to 74). The trend for young adults is in line with the other tests.

Insert Figure 2 about here

If black IQ were constant or falling, the probability of a rise in five comparisons is less than or equal to 1 in 32 (.03). The terminal values are as follows: the WAIS (under 25) terminates at an IQ of 88.08 in 1995, the AFQT at 85.61 in 1997, the SB at 88.40 in 2001, and the WISC at 88.10 in 2002. These give an average of 87.55 and if all trends are projected to 2002, the average rises to 88.2. The average of the median ages is 15.

Differences between black IQs from one standardization to another give estimates of the rate of gain. The WAIS (under 25) shows 3.22 points gained over 17 years (rate = 0.189 points per year). The AFQT shows 3.62 points over 17 years (rate = 0.213). The SB shows 1.79 points over 16 years (rate = 0.112). The WISC shows 1.51 points from 1972 to 1989 (rate = .089) and 4.16 points from 1989 to 2002 (rate = 0.320). Averaging these gives a rate of gain of 0.185 points per year.

The data do not show when recent black IQ gains began. Using 83.5 as the value at the start of gains, the SB trend indicates 1957, the WAIS (under 25) 1971, the WISC

1984, and the AFQT 1987. The SB date is the least plausible in that the earlier gains

began, the more likely that previous scholars would have noticed something.

Insert Table 1 about here

Table 1 pools the data and utilizes proper controls to estimate the black rate of gain

and its standard error.   We cannot reject the hypotheses that gains are the same on all

tests or that gains were constant over time. Therefore, pooling the data to compute a

single rate for all tests for the entire period is appropriate. On the other hand, the

differences between black average scores on different tests and at different ages are

statistically significant. Therefore, the preferred estimates are those with controls for

both age and test.

For blacks under 25, these show a rate of gain of 0.184 IQ points per year. The rate

of gain for those 25 and older in the WAIS data is smaller but our data yield no reliable

estimate for older blacks. The numbers aged 25 to 74 in the WAIS standardization

samples are small. The 95% confidence interval is +/- 0.129. We cannot reject the

hypothesis that older blacks had the same rate of gain as those under 25.

We derived estimates of black IQ gains on whites for each age from 4 to 24. Using

our test/year/age-group data points, and using the regression with controls for test and age

(for those under 25), we projected results before 1987 back to 1972, and results 1987 and

later forward to 2002. To each data point, we subtracted or added the annual rate of gain

(0.184) times the difference between the year the test was administered and the year to

which we were projecting it. We also adjusted each value by subtracting the coefficient

of the corresponding test indicator variable from that administration (with those coefficients normalized to represent the deviation from the average of all tests).

Insert Figure 3 here

In Figure 3, the projected values for 2002 are displayed as squares and those for 1972 as diamonds. An OLS regression line has been fitted to each set of values. Black IQs decline relative to whites with age, indeed, the decline amounts to 11 points between 4 and 24. Despite this, black gains on whites over this 30-year period are close to 5.5 IQ points for all ages below 25.

Figure 3 shows that in 2002, the mean IQ of blacks ranged from 95.5 at age 4 to 84.5 at age 24; and in 1972, from 90 to 79. It puts blacks aged 15 in 2002 at 88.8. Recall that 88.2 was the value suggested by our rough calculations. Figure 3 puts blacks aged 12 at 90.5. We derived our pre-rise estimate of 83.5 from the Coleman Report and the WISC-R and their subjects were 12.5 and 11 respectively, which gives an average age of about 12 (11.75). This would imply that blacks have gained a total of 7 points on whites. But racial differences on the tests used in the Coleman Report may not be comparable to the tests in our sample. It is safer to say that: blacks today aged 12 have a mean IQ of about 90.5; and young blacks have gained 5.52 points on whites over 30 years. Using the 95% confidence limits according to our regression estimates, the gain falls between 4.8 and 6.3 points. It is worth noting that the only data set (the WISC) that covers the entire period of 1972 to 2002 gives a gain of 5.67 points.

Blacks gained on whites even though whites made their own gains. From 1972 to

2002, 12 cases in which the same subjects took both a later and earlier version of a

Wechsler or Stanford-Binet IQ test show an average gain for all Americans of 0.311

points per year (Flynn & Weiss, under review). If both blacks and Hispanics (see

Appendix A) have been gaining at a faster rate, the rate of gain for non-Hispanic whites

(about 75% of the population) would be approximately 0.265. Therefore, the rate of gain

for blacks has been about 0.45 points per year (0.265 + 0.184).

IQ gains and *g* gains

Some attribute the predictive validity of IQ scores to their correlation with the *g* factor.

Whether or not this is true, it raises the question of whether black IQ gains on whites

reflect *g* gains.

To compute *g* scores for the WISC, WAIS, and AFQT (there is no race data for

the SB subtests), we must use subtest scores. To compute the *g* gap between black and

white, we took their average difference on the standardised first principal component of

the subtest correlation matrix multiplied by 15 (thus making the *g* scores equivalent to IQ

scores). By comparing the *g* gap on one test (say the WISC-R) and another test (the

WISC-IV), we can estimate how much the *g*-gap closed. However, our estimates of IQ

trends are based on the age group averages publishers gave us, which makes them non-

comparable. Comparability requires computing IQ differences by summing subtest

differences. These new estimates differ from the old due to our lack of individual data,

that is, our method of aggregating subtests differs from that used by the publishers. The

differences between the two sets of IQ estimates are minimal. Comparability with *g*

gains also entails omitting the correction for the presence of Hispanics in the WISC-R

and WAIS-R standardisation samples from the IQ estimates.

Insert Table 3 about here

Table 3 shows that the estimated *g* gains for blacks average at 93.4% of the

comparable IQ gains, which would imply a *g* gain of 5.16 points from 1972 to 2002.

Yet, when we correlate subtest gains with subtest *g* loadings (the correlation of subtest

scores with *g* factor scores), we get negative values. This means that black gains cannot

be entirely due to changes in *g*. Principal components analysis resolves this conundrum

by suggesting that blacks lost ground on factors other than *g*. We find this unconvincing

and suspect that blacks have gained in some areas and not in others in a pattern unrelated

to both *g* loadings and other factor loadings.   However the trends came about, the brute

fact remains: the standard measure of the *g* gap between black and white declined

virtually in tandem with the IQ gap.

Conclusion

Other scholars provide scores from blacks and whites that took the same test some years

apart or analyze trends (Gottfredson, 2005; Lynn, 1996; Murray, 2005; Vincent, 1991;

Wicherts, 2005). In every case, the samples lack the quality of standardization samples.

All results are compatible with an IQ of 90.5 for black school children in 2002.

However, some show them reaching that value 15 years before our standardisation

samples do, and several show little or no change during the periods they cover (see

Appendix C).   All existing data suggest that since the 1960s, black children have made

large IQ gains relative to whites even if the timing of those gains is uncertain.

The constancy of the black/white IQ gap is a myth. Blacks have gained 5 or 6 IQ

points on whites over the last 30 years. Neither changes in the ancestry of those

classified as black nor changes in those who identify as black can explain more than a

small fraction of this gain. Therefore, environment has been responsible. The last two

decades have seen both positive and negative developments: gains in occupational status

and school funding have been accompanied by more black preschoolers in single-parent

homes and lower income in those homes (Neal, in press).   We believe that further black

environmental progress would engender further black IQ gains.

Appendix A: Data

Insert Table A1 here

Table A1 contains the summary data from the test publishers and the Department of Defence. Break downs by age upon request. The WISC-R and WAIS-R scores for whites have not been corrected for the inclusion of Hispanics (see below). Riverside publishers have requested a note: (1) Controlling for levels of parental education substantially reduces IQ differences between ethnic groups: (2) The SB-5 standardization sample included special education and limited English proficiency groups not included in the SB-4 sample. Our comment: because a higher percentage of blacks than whites are in these categories, this would lower the SB-5 black mean and deflate the SB estimate of black IQ gains.

All white samples were non-Hispanic whites except the 1972 WISC-R and the 1978 WAIS-R samples. Census and Current Population Survey data showed that 5.35 % of the entire population was Hispanic in 1970, 5.13% in 1973, and 5.57% in 1978. The last date corresponded exactly with the WAIS-R. If 5.57% of the total sample were Hispanic, then 6.31% of whites were Hispanic. The percentages for 1970 and 1973 were interpolated to give 5.203% in 1972 at the time of the WISC-R; which meant that 6.15% of whites in its sample counted as Hispanic.

As for Hispanic IQs, the Coleman Report of 1965 (Jensen, 1980, p. 479) showed Hispanics aged 8 to 17 at 12.79 IQ points below non-Hispanic whites; the SB-4 of 1985 (Thorndike, Hagen & Sattler, 1986, pp. 34-36) showed Hispanics aged 2 to 23 at 8.87 points below. These were interpolated to get a value for the WISC-R Hispanics of 11.42

points below non-Hispanic whites. Multiplying 11.42 times .0615 yields the reduction in

the white score due to the inclusion of Hispanics, that is, 0.70 IQ points. The Coleman

Report showed Hispanics aged 17 at 12.975 points below; the SB-4 showed Hispanics

aged 12 to 23 at 8.16 points below. Interpolated these gave a value for the WAIS-R

Hispanics at 9.85 points below. Multiplying 9.85 times .0631 yields the reduction in the

white score due to the inclusion of Hispanics, that is, 0.62 IQ points. Adding the Hispanic

corrections to the white means made little difference. The estimated rate of black IQ

gains rose by less than .03 IQ points per year in all specifications.

Appendix B: Methods

Black IQs for Figure 1 and the AFQT

For the AFQT, we computed average black IQs separately for each age for each

standardization sample. We computed each black subject's age and sample specific IQ

by first finding that subject's percentile ranking among whites of the same age taking the

test in the same year. Percentile ranks were computed using the weights provided us by

the Department of Defence. We then computed a *z*-score from this percentile rank using

the inverse cumulative normal distribution. The *z*-score was multiplied by 15 and added

to 100 to compute the IQ score. Averages for each age, as well as year specific standard

deviations, for blacks were computed (whites were assumed to have an SD of 15 IQ

points by the way we constructed the scores) using the sample weights we were given.

IQs at each percentile of the black distribution for each year were computed using sample

weights and the scores plotted to produce Figure 1.

Average IQs for Figure 2

Average black scores for each test at each point in time were constructed by

computing the weighted average of the average black IQ scores by age provided by the

publishers. For the weights we used the inverses of our estimates of the sampling

variability of the black/white differences for each age group on each test or

$$(A1) \qquad \overline{Q}_{pt}^{\,b} \quad = \quad \sum_{a=1}^{K_p} \frac{\overline{Q}_{pta}^{\,b}}{\hat{\sigma}_{pta}^{2\,Sb}} \Bigg/ \sum_{a=1}^{K_p} \frac{1}{\hat{\sigma}_{pta}^{2\,Sb}}$$

where

$$(A2) \qquad \overline{Q}_{pta}^{b} \quad = \quad 100 - \frac{15(\overline{S}_{pta}^{b} - \overline{S}_{pta}^{w})}{\hat{\sigma}_{pt}^{w}}$$

(except for the AFQT where $\overline{Q}_{pta}^{b}$ was just the weighted average black IQ for that age

category since those scores were computed as differences from non-Hispanic white

scores and white *SD* was assumed to be 15),

$$(A3) \qquad \hat{\sigma}_{pta}^{2\,Sb} \quad = \quad \left(\frac{15}{\hat{\sigma}_{pt}^{w}}\right)^{2} \left(\frac{\hat{\sigma}_{pt}^{2\,w}}{N_{pta}^{w}} + \frac{\hat{\sigma}_{pt}^{2\,b}}{N_{pta}^{b}}\right),$$

$$(A4) \qquad \hat{\sigma}_{pt}^{w} \quad = \quad \sqrt{\sum_{a=1}^{K_{p}} N_{pta}^{w} \hat{\sigma}_{pta}^{2\,w} \bigg/ N_{pt}^{w}}$$

(except for the AFQT where the population white standard deviation is 15 by

construction and that is used instead of the sample standard deviation), K_{p} is the number

of age groups results for test p were broken down into by the publishers who provided the

data, $\overline{S}_{pta}^{b}$ is the average reported score of blacks on test p at time t for age group a,

$\overline{S}_{pta}^{w}$ the average in that group for whites, N_{pta}^{b}, N_{pta}^{w} the black and white age specific

sample sizes respectively, $\hat{\sigma}_{pta}^{2\,w}, \hat{\sigma}_{pta}^{2\,b}$ the white and black age specific sample variances,

and N_{pt}^{w} the total number of whites taking test p at time t. Our estimates of the sampling

variability are not exact as all means reported to us were computed using sample weights.

Our sources did not provide us with adequate information to get exact estimates of the

sampling variability of the age group means we use as data points. However, our

estimated means are still consistent and should be more efficient than unweighted estimates.

Regression Estimates for Figure 3 and Table 1

The rates of gain for black IQ reported in Table 1 come from a series of regressions of black IQ scores on different sets of independent variables. The observations for these regressions were the 69 test-year-age specific black means computed as just described. Independent variables were the year the test was administered, dummy variables for the tests (SB, WAIS, and WISC with the AFQT as the left out category), and average age of subjects, average age squared, and average age cubed. The rate of gain reported is the coefficient of the year variable. Generalized Least Squares (*GLS*) was used to allow for heteroskedacticity using the weighting procedure suggested by Dickens (1990). We also computed White robust standard errors to allow for general correlation of errors across age groups within each administration of each test. This allowed for correlation in black means across age groups due to either the norming procedures used by the publishers or our use of a common white standard deviation computed from all age groups to normalize black/white differences. Whenever robust standard errors and GLS standard errors differed, we reported the larger value.

Figure 3 was constructed by taking the 69 age-year-test specific estimates of black IQ and projecting them either forward to 2002 or back to 1972 using the regression of black IQ on year, test dummies, age, age squared, and age cubed for those under 25. The procedure is described in the text.

Appendix C: Trends from other samples

Vincent (1991) presents results on Raven's Progressive Matrices. The 1973

sample was 380 subjects from a rehabilitation unit in Houston (Vincent & Cox, 1974).

At least one race had more education than the group it represents. The 1985 sample

consisted of 631 whites and 209 blacks from Decatur, Alabama (Raven, 2000, pp. 19-21).

Between 1973 and 1985, black IQ (normed on whites) goes from 84 to 93, a huge gain.

However, the first sample had a median age of 29, the second 9. By our estimates, blacks

lose about 10 points on whites between those ages (the rate slows down after age 24).

Therefore, the 1973 mean must be raised to 94 (84 +10).

Gottfredson (2005) concludes that the black/white gap is between 0.8 and 1.2

SDs. She lacks our most recent data and her estimates do not prelude our 2002 value of

0.63 SDs. For example, she refers to the Wonderlic Personnel Test, a 12-minute test with

50 items. Between 1970 and 2001, that test was normed four times on samples of job

applicants. Increasing reluctance to record race and age reduced whites from 123,000 to

15,600 and blacks from 34,000 to 2933. Long (2006) notes that more Americans aged 16

to 24 remain in education (rather than seeking work) today than in the past, and that more

older workers of high quality are made redundant, changes that affect the races

unequally.

Wonderlic (2006) provides data for 1970, 1983, 1992, and 2001. Blacks aged 16

to 24 begin at an IQ of 84.6 in 1970, rise to 87 in 1992, and remain there in 2001. Blacks

of all ages (16 and above) begin at 85 in 1970 and are again at 85 in 2001. An IQ of 87

for young blacks (median age 20) is only one point above the 86 derived from

standardization samples. But rather than a gradual rise between 1970 and 2001, there is a sudden rise in 1992 with no change in any other year.

Murray (2005) notes that standardizations of the Kaufman-ABC in 1983 and 2004 show black IQs of 93 and 92.1. The K-ABC subtests were selected to minimize the black/white gap (Jensen, 1984) and diminish *g* in favour of short-term memory (Naglieri and Jensen, 1987). Jensen argues that the 1983 sample contained too great a range of ability, thus yielding an inflated SD and a diminished black/white gap (measured in SD units).

Standardization samples of the Woodcock-Johnson (WJ) are excellent but the race data does not come from them. A research sample is taken from the full sample and a sub-sample of that (those who took all subtests used to compute IQ) used to calculate *g* scores. This sub-sub-sample comprised 90% of the 1987 research sample and just over 50% of the 1999 sample.   Racial comparisons are based on blacks and whites in the sub-sub-sample with blacks more likely to be absent in both years.   Wicherts (2005) gave us data for ages 1 to 65: black IQ was steady at 88 between 1987 and 1999; for those under 25, it stood at 90 in 1987 and 88.70 in 1999.

Our analysis of standardization samples showed black IQ (age 12) rising from 85 in 1972 to 90.5 in 2002. For young blacks, the K-ABC gives 93 in 1983, Raven's 93 in 1985, the WJ 90 in 1987, and the Wonderlic 87 in 1992. Averaging these puts black IQ at 91 circa 1987. The three tests with results both near and after that date show no gains thereafter.   The imperfect data support the contention that black school children attained

an IQ of 90.5 by 2002, but has them reaching that value 15 years earlier than our results

from standardisation samples.

Lynn (1996, p. 272) uses results by age to infer trends. Age patterns do not chart

trends over time but rather, reflect an altering black/white gap as cohorts age. His 85.83

(our convention) for blacks aged 6 to 17 (1986) is close to our 84 for the WISC-III

(1989).

References

Department of Defense (2005). Unpublished standardization data for the AFQT from the

Profiles of American Youth of 1980 and 1997 provided by the US Department of

Defense.

Dickens, W. T. (1990). Error components in grouped data: Is it ever worth weighting?

Review of Economics and Statistics, 72, 328-333.

Farley, R. (1995). *State of the union: America in the 1990s*. New York: Russell Sage

Foundation.

Flynn, J. R. (1987). Flynn replies to Nichols. In S. Modgil & C. Modgil (eds.), *Arthur*

Jensen: Consensus and controversy (pp. 234-235). Lewes, Sussex: Falmer Press.

Flynn, J. R., & Weiss, Lawrence G. (under review). American IQ gains from 1932 to

2002: The significance of the WISC subtests.

Gottfredson, L. S. (2005). Implications of cognitive differences for schooling within

diverse societies. In C. L. Frisby & C. R. Reynolds (eds.), *Comprehensive*

handbook of multicultural school psychology (pp. 517-555). New York: Wiley.

Harcourt (2005a). Unpublished standardization data for Wechsler tests: the WAIS-R,

WAIS-III, WISC-R, WISC-III, and WISC-IV. Copyright © 2005 by Harcourt

Assessment, Inc. Used with permission. All rights reserved.

Harcourt (2005b). Unpublished data on factor analysis of Wechsler tests: the WISC-III

and WISC-IV. Copyright © 2005 by Harcourt Assessment, Inc. Used with

permission. All rights reserved.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure*

in American life New York: The Free Press.

Jensen, A. R. (1980). *Bias in mental testing*. London: Methuen.

Jensen, A. R. (1984). The black-white difference on the K-ABC: Implications for future

tests. *Journal of Special Education, 18,* 377-408.

Jensen, A .R. (1987). The *g* beyond factor analysis. In R. R. Ronning, J. A. Glover, J. C.

Conoley, & J. C. Witt (Eds.), *The influence of cognitive psychology on testing*

(pp. 87-142). Hillsdale, NJ: Erlbaum.

Jensen, A. R. (1992). Spearman's hypothesis: Methodology and evidence. *Multivariate*

Behavioral Research, 27, 225-233.

Long (2006).   Personal communication from Eliot R. Long, Research Manager, Data

Research Services, Wonderlic, Inc., January 5, 2006.

Lynn, R. (1996). Racial and ethnic differences in intelligence in the United States on the

Differential Ability Scale. *Personality and Individual Differences, 20*, 271-273.

Murray, C. (2005). The inequality taboo. *Commentary*, September, 2005, 13-22. See:

http://www.commentarymagazine.com/production/files/murray0905.html (fn 41

and 44).

Naglieri, J. A., & Jensen, A. R. (1987). Comparison of black-white differences on the

WISC-R and the K-ABC: Spearman's hypothesis. *Intelligence, 11,* 21-43.

Neal, D. (in press). Why has black-white skill convergence stopped? In *Handbook of*

Economics of Education, edited by E. Hanushek and F. Welch. New York:

Elsevier.

Raven, J. (2000). *Manual for Raven's Progressive Matrices and Vocabulary Scales:*

Research supplement no. 3. Oxford: Oxford Psychologists Press.

Riverside (2005). Unpublished standardization data for the SB-5. Copyright © 2005 by Riverside Publishers. Used with permission. All rights reserved.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales*, Fifth Edition, Technical Manual. Itasca, IL: Riverside Publishing.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law 11*, 235-294.

Segall, D. O. (1997). Equating the Cat-asvab. In Sands, W. A., Waters, B. K., & McBride, J. R. (eds*.), Computerized adaptive testing; From inquiry to operation* (pp. 181-198). Washington, D.C.: American Psychological Association.

Staples, R. (1985). Changes in black family structure: The conflict between family ideology and structural conditions. *Journal of Marriage and Family, 47,* 1005-1013.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *The Stanford-Binet Intelligence Scale: Fourth Edition -- Technical Manual.* Chicago: Riverside Publishing.

Vincent, K. R. (1991). Black/white IQ differences: does age make the difference? *Journal of Clinical Psychology*, *74,* 266-270.

Vincent, K. R., & Cox, J. A. (1974). A re-evaluation of Raven's Standard Progressive Matrices. *Journal of Psychology, 88,* 299-303.

Wicherts (2005). Comparison of white and black groups in the US research

samples of the Woodcock-Johnson-R and the Woodcock-Johnson-III.

Unpublished data provided by J.M.Wicherts, December 19, 2005.

Wonderlic (2006).   Norms by race and age for the Wonderlic Personnel Test N:

Normative studies 1970, 1983, 1992, & 2001. Unpublished data: Copyright ©

2006 by Wonderlic, Inc. Used with permission. All rights reserved.

TABLE 1

Annual Rate of Black Gain in IQ Points

	Controls			
Sample	None	Age[1]	Test	Test and Age[1]
All Ages	.195	.187	.183	.188
	(.046)	(.025)	$(.031)^2$	$(.021)^2$
Under 25	.212	.161	.190	.184
	(.057)	$(.033)^2$	$(.033)^2$	$(.025)^2$

Note: Rates of gain are the coefficient of a year variable in a regression of black IQs by age, test, and year on year and other controls specified. The specifications of the regressions are described in the methods appendix. Standard errors are the maximum of those from the generalized least squares procedure or White robust standard errors.

1. Controls for age are average age of group in years, average age squared, and average   age cubed. When all ages are included a separate age polynomial is estimated for those over 24.

2. GLS standard errors (others are White robust)

Table 2

Change in g vs. change in Full Scale IQ

(in IQ points)

	WAIS R to III		WISC
	full sample	<25	R to IV
g gain	1.17	2.57	4.67
Full Scale IQ gain (from sub-tests)	1.20	2.82	4.93
Full Scale IQ gain (from publisher)	1.09	2.60	4.96
Average correlation of W-B difference with g loadings[1]	.65	.74	.86
Average correlation of subtest gains with g loadings[1]	-.28	-.73	-.38

1. Common subtests only

Table A1						
IQ Means and Standard Deviations for Whites and Blacks						
Test		Mean IQ		Standard Deviation		Number of Observations	
		White	Black	White	Black	White	Black
Stanford-Binet	4	103.6	90.0	15.37	13.86	3691	711
	5	102.9	92.1	13.93	14.47	2070	384
WISC	R	102.3	86.4	14.08	12.63	1870	305
	III	103.5	88.6	13.86	12.83	1543	337
	IV	103.2	91.7	14.52	15.73	1402	343
WAIS	R	101.4	86.8	14.65	13.14	1664	192
	III	102.6	89.1	14.81	13.31	1523	247
WAIS < 25	R	101.2	87.0	14.28	13.54	519	72
	III	102.6	90.9	14.59	12.31	413	93
AFQT	80	100.0[1]	82.0	15.00[2]	13.63	5533	2298
	97	100.0[1]	85.6	15.00[2]	13.23	2880	1191

1. This is 100 by construction. See text.
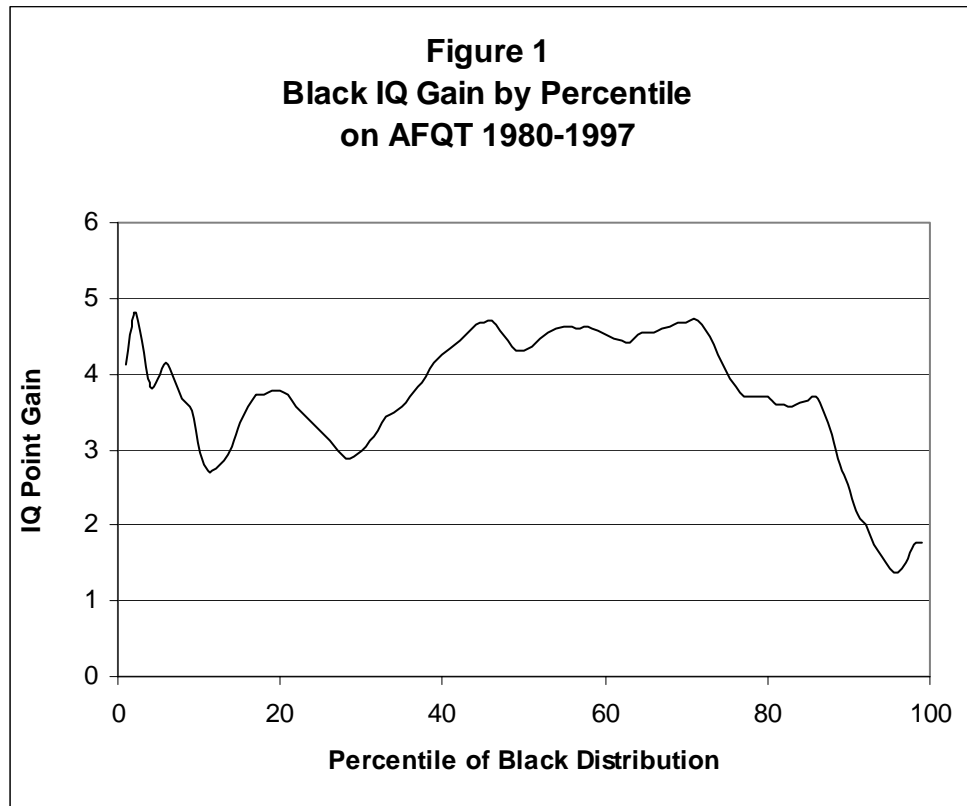2. This is 15 by construction. See text.

Figure 1
Black IQ Gain by Percentile
on AFQT 1980-1997

Figure 2
Black Scores on Four Tests
of Cognitive Ability
(white average = 100)

Figure 3
Projected Black IQ by Age
in 1972 and 2002