

Changes over time in the black–white difference on mental tests: Evidence from the children of the 1979 cohort of the National Longitudinal Survey of Youth

Charles Murray

American Enterprise Institute, 1150 Seventeenth St., NW Washington, DC 20036, USA

Received 4 May 2006; received in revised form 19 June 2006; accepted 12 July 2006
Available online 24 August 2006

Abstract

Data for three Peabody achievement tests and for the Peabody picture vocabulary test administered to children of women in the 1979 cohort of the National Longitudinal Survey of Youth show that the black–white difference did not diminish for this sample of children born from the mid 1970s through the mid 1990s. This finding persists after entering covariates for the child’s age and family background variables. It is robust across alternative samples and specifications of the model. The analysis supplements other evidence that shows no narrowing of the black–white difference in academic achievement tests since the late 1980s and is inconsistent with recent evidence that narrowing occurred in IQ standardizations during the same period. A hypothesis for reconciling this inconsistency is proposed.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Intelligence; IQ; Black-white differences; Peabody Test Battery; National Longitudinal Survey of Youth; NLSY; Longitudinal test data; Achievement tests, Race differences

1. Introduction

Longitudinal trends in the black–white (B–W) difference in mental test scores have both policy and theoretical implications. The policy implications arise because the assumption that black and white test scores will converge over time is woven into the rationale for affirmative action, most explicitly in the Supreme Court’s 2003 decision regarding the University of Michigan Law School. The Court observed that affirmative action must be time limited if it is to be constitutional, but cited rising numbers of African-American applicants with high test scores as evidence that race preferences could in fact be time limited, concluding that “we expect that 25 years

from now the use of racial preferences will no longer be necessary” (*Grutter v. Bollinger* 539 U.S. 306 [2003]).

The theoretical implications involve the debate over environment and genes as causes of the B–W difference. If environment is the exclusive cause, continued improvement in the economic and educational status of African-Americans should be associated with continued narrowing of the B–W difference; insofar as genes play a role, that narrowing has limits.

The empirical evidence is consistent for tests intended to measure academic achievement: reduction in the B–W difference occurred during the 1970s and into the 1980s, then stopped. Over the period from 1965 to 1992, *Hedges and Nowell (1998)* found a reduction from 1.18 to .82 standard deviations (S.D.s) in academic achievement tests administered to 12th graders. Data

E-mail address: charlesmurray@adelphia.net.

from the National Assessment of Educational Progress (NAEP) show reductions in the B–W difference in both math and reading scores for students aged 9, 13, and 17, from 1972 to the end of the 1980s, in magnitudes ranging from .25 to .68 S.D.s (Perie & Moran, 2005, supplemented with S.D.s provided by personal communication from the National Center for Educational Statistics). The College Board's report on profiles for college-bound seniors (College Board, annual) reveals that the B–W differences on the verbal and math tests dropped by .37 and .35 S.D.s respectively from 1972 to 1991 (author's analysis), and there is reason to believe this represents a real convergence in scores, not merely changes in the composition of the pool of SAT test-takers (Herrnstein & Murray, 1994).¹

The B–W difference on achievement tests did not continue to narrow in the 1990s. The closest convergence on the NAEP math and reading tests was reached in the 1986, 1988, or 1990 test waves, depending on the test and age group. In the SAT, the B–W difference has widened since 1991, from .87 to .98 S.D.s on the 2005 verbal test and from .91 to 1.03 S.D.s on the 2005 math test.

Regarding tests specifically designed to measure cognitive ability rather than academic achievement, a vigorous debate is underway. In two recent literature reviews, Gottfredson (2005) and Rushton and Jensen (2005) examined test results as far back as World War I and found no evidence of a trend in the IQ difference. A meta-analysis of the B–W difference on cognitive tests (Roth, Bevier, Bobko, Switzer, & Tyler, 2001) did not analyze trends, but their conclusion from test results extending into the 1990s was that the B–W difference for highly *g* loaded test batteries centers on 1.1 standard deviations, in line with characterizations of the historic B–W difference (Herrnstein & Murray, 1994; Rushton & Jensen 2005; Dickens & Flynn, in press-a).

On the other side, Vincent (1991) presented evidence for a narrowing difference in IQ scores from the 1980s, but it was vulnerable to technical criticism (Jensen, 1998). A stronger case has been made for the vocabulary test administered by the General Social Survey. Lynn (1998) examined the B–W difference by test year from 1974 to 1996 and found a small reduction that did not reach statistical significance. Huang and Hauser (2001) reana-

lyzed the data by birth cohort and found a highly significant reduction. Most recently, Dickens and Flynn (in press-a) presented evidence from the standardization samples for the Wechsler scales for children and adults (WISC and WAIS), the Stanford–Binet, and the Armed Forces Qualification Test (AFQT), pointing to a reduction of 3 to 6 points in the B–W difference between 1972 and 2002. A reply (Rushton & Jensen, in press) presented contemporaneous test results that showed no reduction and argued that the totality of the evidence failed to support their conclusion. This position was disputed in turn in a rejoinder (Dickens & Flynn, in press-b).

The purpose of this article is to bring new data to bear on the issue. The children of women in the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79) offer a database with six desirable characteristics for this purpose: It is longitudinal, has large samples, tracks down children regardless of whether they were attending school, includes both achievement and cognitive tests, provides results over time using the same instruments, and has extensive family background data, including a measure of maternal IQ.

2. Method

2.1. Sample

The NLSY79 consists of a nationally representative sample augmented by oversamples of blacks, Hispanics, and low-income whites. The analyses that follow are based on children of mothers whose ethnicity was classified as black or non-Hispanic white, and to tests administered when the subject's age was appropriate to kindergarten through the 12th grade. I will hereafter refer to the children as "black" or "white" based on the mother's ethnicity.

Beginning in 1986 and every two years thereafter, children of the NLSY79 women have been given three subtests of the Peabody Individual Achievement Test battery: reading recognition (PIAT-RR), reading comprehension (PIAT-RC), and mathematics (PIAT-M). They have also been given the revised Peabody picture vocabulary test (PPVT-R). Reliabilities and validities for these widely used tests, presented in Dunn and Markwardt (1970) and Robertson and Eisenberg (1981), are competitive with comparable tests for school-age children (Baker, Keck, Mott, & Quinlan, 1993).

The nature of the mental skills measured by these four tests varies. The items in the reading recognition test measure skills such as matching letters, naming names, and reading single words aloud. It is a basic test of literacy, with minimal interpretive requirements. In contrast, the items in the reading comprehension test

¹ S.D.s were not available for the 1972 SAT data. I used the S.D.s from the first year when they were reported, 1980. Differences expressed in S.D.s in the text are always based on the within group S.D.s,

$$\frac{(\bar{X}_a - \bar{X}_b)}{\sqrt{(N_a\sigma_a^2 + N_b\sigma_b^2)/(N_a + N_b)}}$$

where *N* is the sample size, *X* is the sample mean, σ is the standard deviation, and the subscripts a and b denote each group.

ask the child to read a sentence silently, then select from four pictures the one which best portrays the meaning of the sentence. It thus measures both reading skills acquired through schooling and cognitive ability to derive meaning from combinations of words. Similarly, items in the mathematics test are multiple-choice problems that test both learned mathematical knowledge and cognitive ability to reason mathematically.

Unlike the three achievement tests, the PPVT-R does not require any trained reading or calculation skills. It asks the child to respond to a spoken word by selecting the one of four pictures that best describes its meaning. The PPVT-R is a widely used screening test for verbal IQ, based on a phenomenon identified early in the development of IQ tests: the number of words a person knows does not primarily reflect exposure to words as determined by schooling and socioeconomic status, but is “at once a measure of his learning ability, his fund of verbal information, and of the general range of his ideas” (Matarazzo, 1972: 218). Empirically, the score on the vocabulary subtest in an IQ test battery typically has the highest correlation with full-scale IQ of any subtest (Jensen, 1980). The Peabody picture vocabulary test specifically has shown median correlations of .62 and .64 with full-scale IQ as measured by the Stanford–Binet and the WISC respectively (Robertson & Eisenberg, 1981).

All children who could be located and who were eligible for a test or retest in a given year were included in each test wave. Through 2002, the last test wave included in the data for this analysis, the maximum number of administrations of the same test to the same child was five. The median was two administrations of each test per child. The database is arranged as one observation per child per test year. The sample for analysis includes all children who had test scores taken at ages appropriate to grades K–12, and who had complete data on birth date, race, age-appropriate grade in school, and mother’s AFQT score. This produced 21,125 observations generated by 6209 children, of whom 3652 were non-Hispanic white and 2557 were black.

2.2. Variables

2.2.1. PIAT-RR, PIAT-RC, PIAT-M, and PPVT-R

The tests scores are expressed in terms of the national norm samples for the tests, each of which had a mean of 100 and an S.D. of 15. The NLSY manual notes that the versions of the achievement tests used for the NLSY79 children were normed in the late 1960s and by now produce scores that are misleadingly high if they are interpreted relative to a national mean of 100 (Baker et

al., 1993). The three achievement tests have a range of 65–135, while the PPVT-R has a range of 40–160. All scores are normed for age.

2.2.2. Birth date

The variable for estimating changes in the B–W difference over time is the birth date, expressed as a continuous variable starting with last two digits of the year. The earliest birth occurred at the end of 1970 (70.96), but 97% of the children were born from 1975 onward. The effective range for birth date is best understood as the mid 1970s through the mid 1990s. The highest value for birth date is 97.63.

2.2.3. Age of the child

The variable for estimating the effects of the child’s age is the subject’s age-appropriate grade in school during the test year, expressed in 1-year increments from 0 (kindergarten) through 12. Grade was chosen rather than actual age because students born in the last months of a calendar year are likely to be a grade behind those born earlier, an important consideration in assessing achievement test scores. “Age-appropriate” means that a child who has repeated a grade or dropped out is coded according to the grade to which his age corresponds, not his actual grade in school. Coding for the age-appropriate grade variable assumes that all children born before September were eligible for kindergarten 5 years after the calendar year of birth, and those born from September through December were eligible 6 years from the calendar year of birth.

2.2.4. Mother's cognitive ability

The measure of the mother’s cognitive ability is her score on the Armed Forces Qualification Test (AFQT), one of the most highly *g*-loaded paper-and-pencil cognitive tests (Herrnstein & Murray, 1994), normalized by age and expressed in an IQ metric ranging from 55 through 145.

2.2.5. Mother's education

The measure of mother’s education is her completed years of education as of the year of the child’s birth, ranging from 0 through 20.

2.2.6. Family income

Income is represented by the median total family income during the 5 years from the child’s birth until age 5, converted to constant 2002 dollars, with a range from zero through \$217,075. Since the role of income is presumptively nonlinear (e.g., the difference between \$10,000 and \$30,000 in family income can be expected to have greater effects on the home environment than the difference between \$100,000 and \$120,000), the income

variable, expressed in thousands of dollars, was logged, with all incomes less than \$1,000 assigned a value of zero.

2.2.7. Family structure

Family structure is represented by a dummy variable coded 1 if the mother was married when the child was born and coded 0 if she was not. More detailed breakdowns of family structure from birth through age 4 were examined, but did not add to the information conveyed by this binary characterization.

2.2.8. Mother's age at birth

The chief technical problem in drawing conclusions from the NLSY data is that the independent variable of central interest, the child's birth date, is highly correlated with mother's age at birth, because all of the NLSY79 mothers were born within a comparatively narrow period from 1957 through 1964. Thus, for example, none of the children born in 1975 could have a mother

older than 19, while none born in 1995 could have a mother younger than 30. The bivariate correlation between mother's age at birth and child's birth date in the sample used for analysis is .90. Further, mother's age at birth has a nontrivial bivariate correlation with test scores, ranging from .17 for the PIAT-RC to .25 for the PIAT-M, and mean age at birth varied substantially by race, with means of 25.2 years for whites and 23.1 years for blacks. Together, these issues pose a threat that any observed relationship between the child's birth date and test scores is contaminated by multicollinearity.

To deal with this problem, the data were analyzed using three different treatments of the birth-date/maternal-age relationship. Model 1 enters mother's age at birth as a continuous variable. Model 2 treats mother's age at birth as a vector of dummy variables consisting of 4-year segments beginning with 12–15 and ending with 36–39.

Model 3 used matching by maternal age at birth to eliminate any possibility that race differences in age at

Table 1
Descriptive statistics for the variables in the models

Variables by sample	White				Black			
	Mean ^a	S.D.	Subjects	Observations	Mean ^a	S.D.	Subjects	Observations
Full sample								
<i>Dependent variables</i>								
PIAT-RR (reading recognition)	106.98	13.96	3628	11,597	99.52	14.77	2543	8966
PIAT-RC (reading comprehension)	104.76	13.03	3337	9615	96.58	13.90	2455	7655
PIAT-M (mathematics)	104.34	13.05	3632	11,661	94.86	13.19	2548	9002
PPVT-R (picture vocabulary test)	99.88	16.01	3171	5453	81.84	17.05	2324	4446
<i>Independent variables</i>								
Child's birth date (year, last 2 digits)	85.78	4.87	3652	11,925	84.19	4.86	2557	9200
Child's age-appropriate grade	3.87	2.76	3652	11,925	4.39	2.89	2557	9200
Mother's AFQT score	100.38	12.99	3652	11,925	84.57	10.98	2557	9200
Mother's years of education at birth	12.80	2.18	3290	10,948	12.02	1.88	2141	7833
Median family income from birth to age 4 (\$000)	51.03	34.70	3589	11,797	27.31	23.53	2474	8982
Mother unmarried at birth (%)	12.4%		3111	10,719	62.9%		2249	8226
Mother's age at birth (years)	25.15	4.63	3652	11,925	23.15	4.61	2557	9200
Matched sample								
<i>Dependent variables</i>								
PIAT-RR (reading recognition)	106.92	13.66	1946	2796	100.33	14.87	1620	2832
PIAT-RC (reading comprehension)	104.29	12.66	1689	2343	97.33	13.90	1461	2405
PIAT-M (mathematics)	104.31	12.97	1954	2814	95.72	13.64	1622	2855
PPVT-R (picture vocabulary test)	100.27	15.02	2011	2916	82.81	17.27	1653	2916
<i>Independent variables</i>								
Child's birth date (year, last 2 digits)	85.03	4.08	2011	2916	85.17	3.92	1653	2916
Child's age-appropriate grade	3.91	2.45	2011	2916	3.96	2.41	1653	2916
Mother's AFQT score	99.90	12.29	2011	2916	85.40	11.13	1653	2916
Mother's years of education at birth	12.35	1.99	2011	2916	12.01	1.91	1653	2916
Median family income from birth to age 4 (\$000)	46.28	30.51	2011	2916	28.31	24.39	1653	2916
Mother unmarried at birth (%)	12.0%		2011	2916	59.1%		1653	2916
Mother's age at birth (years)	23.99	4.01	2011	2916	23.98	4.02	1653	2916

^a Means are computed from observations rather than subjects.

birth accounted for the observed trends in the B–W difference. The matching process produced a subsample of 2,916 observations for children of each race. The means and distributions for mother’s age at birth for the white and black samples were effectively identical (means and S.D.s for whites and blacks were both within .01 years of each other). Model 3 was implemented by entering mother’s age at birth both as a continuous variable (Model 3a) and as a vector of dummy variables grouped in 4-year segments (Model 3b).

Descriptive statistics for the variables employed in the analysis are shown in Table 1.

2.3. Analytic procedure

The analysis consists of a random effects generalized least squares (GLS) analysis in which the group variable (i.e., the unit comprised of repeated measures) is the child and the panel variable is test year. Robust standard errors were obtained, adjusted for repeated measures of the same child. For the analysis of trends over time for each race, each specification of the model was implemented separately for each race, with the child’s birth date as the variable of interest. For the analysis of changes in the B–W difference over time, the independent variables were entered along with a categorical variable for race

(black=1) and interaction terms by race for each of the other independent variables, with the interaction term between race and birth date being the variable representing the net change in the B–W difference. A positive coefficient indicates that black scores rose with birth date after taking all the other covariates into account, thereby signifying a narrowing B–W difference; a negative coefficient indicates that black scores fell with birth date, signifying a widening B–W difference.

3. Results

3.1. Trends by birth date by race

Fig. 1 below shows the trends for each of the four tests by birth date, unadjusted for any covariates. Each child in the sample is represented by a single score (the median score for children tested more than once). The trend lines are produced by a bivariate regression of test score on child’s birth date. The squares and circles show means by birth year when the sample size for that year was 50 or more.

Unadjusted test scores rose over time for both blacks and whites for all of the tests. As an inspection of the regression lines indicates, the B–W difference remained unchanged for the two reading tests and increased slightly for mathematics and the PPVT-R.

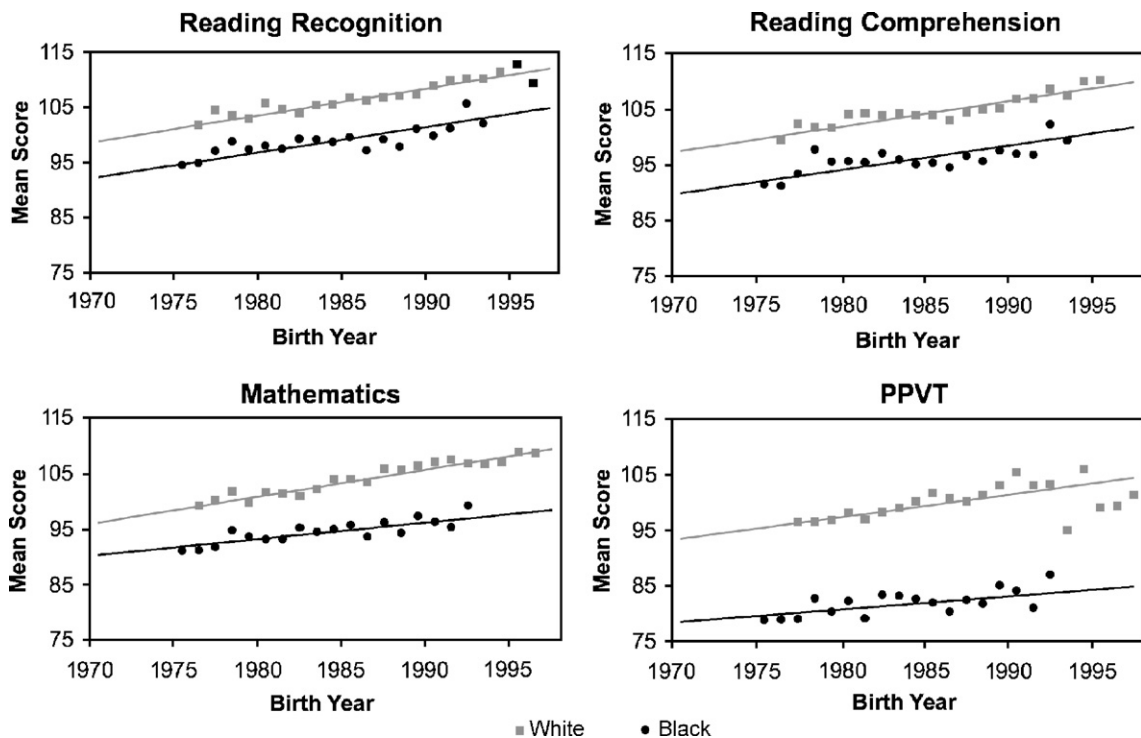


Fig. 1. Unadjusted trends in test score by birth date.

Table 2
Change in white and black test scores per year of child's birth date

Race and model	<i>b</i> for child's birth date			
	PIAT-RR	PIAT-RC	PIAT-M	PPVT-R
Whites				
Model 1	-0.06	-0.20*	0.02	-0.11
Model 2	-0.10	-0.22**	0.02	-0.12
Model 3a	-0.14	-0.36**	0.00	-0.09
Model 3b	-0.16	-0.33**	0.06	-0.11
Blacks				
Model 1	-0.06	-0.33**	-0.12	-0.31
Model 2	-0.01	-0.24*	-0.10	-0.39*
Model 3a	-0.10	-0.51***	-0.19	-0.31
Model 3b	-0.07	-0.44***	-0.18	-0.40*

*** denotes $p < .001$, ** denotes $p < .01$, * denotes $p < .05$.

Variables: Child's birth date, child's age-appropriate grade, family income, and mother's IQ, education, marital status, and age at birth.

The magnitudes of the B–W difference on the fitted values for the three PIAT tests range from 6.5 to 11.0 points depending on test and birth year (about 0.5 to 0.9 S.D.s). Differences of this magnitude are characteristic of the differences on subtests that have yielded a B–W difference averaging 15–16 points for a full test battery (Herrnstein & Murray, 1994). The B–W differences on the fitted values for the PPVT-R were much larger, ranging from 15.7 to 19.7 points, but it should be noted that the standard deviation for the PPVT-R scores among the NLSY children was also unusually large (18.2 points for the median scores used in Fig. 1). Table 2 shows what happens to the unadjusted results when the covariates are entered in the random effects GLS analysis, focusing on the variable of central interest, child's birth date. The full results for the regression are shown in Appendix Table A-1.

Once the covariates have been taken into account, the rise in scores shown in Fig. 1 disappears. Instead, the overall story is one of flat or falling scores for both white and black children over time. The trends over time varied by race, as follows:

Whites: After adjusting for the covariates, the overall trend in white test scores over time was markedly down for reading comprehension, reaching statistical significance in three of the four models. The trend was effectively flat for math and slightly but insignificantly downward for reading recognition and the PPVT-R.

Blacks: Black scores fell on all four tests after adjusting for the covariates, in all four specifications of the model. The drop in reading comprehension reached statistical significance in all four models. The drop in the PPVT-R was of similar magnitude in all four models (.31–.40 points per birth year) and reached statistical significance in both of the models using dummy variables for mother's

age at birth. The reductions in reading recognition and math were smaller and did not reach statistical significance in any of the specifications of the model.

The role of other independent variables: This analysis was not designed to explore the role of the other independent variables in detail, but these generalizations apply:

The role of the child's age at time of testing, represented by age-appropriate grade in Table 2, was strikingly different for white and black children. For whites, scores in math and the PPVT-R rose significantly with age, while falling in reading comprehension. For blacks, scores fell with age on all the tests except the PPVT-R. These results were substantial for the full sample but ambiguous in the sample matched on mother's age at birth.

The coefficients for mother's AFQT score were large and statistically significant at the .001 level for all four tests in all of four specifications of the model for both whites and blacks.

The mother's education at the time the child was born and family income from birth through age four were often but not always significant independently of her AFQT score. No clear pattern characterizes these results.

The mother's marital status at birth did not have a significant relationship to test scores independently of the other variables in any of the tests for either whites or blacks.

The variable that poses the main threat to the validity of the analysis, mother's age at birth, was not statistically significant. This is a function of the other covariates that capture some of the reasons that age at birth is important—the mother's IQ, years of education, and the family's income, all of which are all positively correlated with mother's age at birth (the Pearson r s are .28, .50, and .38 respectively). When the analyses are replicated dropping those variables, mother's age at birth is always significant at the .002 level or higher.

It is this cluster of effects—the children born later in time were, on average, born to smarter, better-educated, more affluent women, as well as to women who were older and presumptively more mature—that accounts for the rising trend lines by birth date shown in Fig. 1 and the reversal of those slopes in the multivariate analyses.

3.2. Trends in the B–W difference over time

Full results for the analysis using a dummy variable for race and interaction terms are shown in Table 3 on the next page. Table 4 below summarizes the results for the interaction term between race and birth date, the variable that estimates the change in the B–W difference. Table 4 also recasts these coefficients in

Table 3
Results of the GLS analysis with interaction terms

Variables by model	PIAT-RR		PIAT-RC		PIAT-M		PPVT-R	
	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.
<i>Model 1: Mother's age at birth as a continuous variable</i>								
Child's birth date	-0.06	0.10	-0.20	0.09*	0.02	0.09	-0.11	0.13
Child's age-appropriate grade	0.01	0.04	-1.03	0.05***	0.26	0.04***	0.49	0.07***
Mother's AFQT score	0.26	0.02***	0.26	0.02***	0.26	0.02***	0.34	0.03***
Mother's years of education at birth	0.42	0.13***	0.22	0.13	0.53	0.12***	0.84	0.19***
Median family income from birth to age 4, logged	1.98	0.36***	1.32	0.35***	1.61	0.33***	0.68	0.51
Mother unmarried at birth (1=yes)	0.14	0.74	-0.78	0.69	0.33	0.64	-1.34	0.94
Mother's age at birth	0.11	0.11	0.21	0.10*	0.07	0.09	-0.06	0.14
Race (black=1)	-7.28	10.72	1.31	9.69	5.06	9.61	-25.20	14.80
Child's birth date × race	0.00	0.16	-0.13	0.14	-0.14	0.14	-0.20	0.22
Child's age-appropriate grade × race	-0.91	0.07***	-1.03	0.07***	-0.36	0.06***	0.24	0.11*
Mother's AFQT score × race	0.08	0.04*	0.09	0.03**	0.04	0.03	0.17	0.05***
Mother's years of education at birth × race	0.48	0.24	0.40	0.21	0.16	0.21	0.71	0.32*
Median family income × race	-0.82	0.55	-0.38	0.51	-0.86	0.49	0.29	0.77
Mother unmarried at birth (1=yes) × race	-0.19	0.95	0.57	0.87	-0.91	0.84	1.74	1.24
Mother's age at birth × race	-0.03	0.17	0.04	0.15	0.08	0.15	0.23	0.23
Constant	70.36	6.45	87.76	6.12	61.38	5.84	62.46	8.90
<i>R</i> ² within	0.04		0.19		0.00		0.03	
<i>R</i> ² between	0.25		0.32		0.31		0.39	
<i>R</i> ² overall	0.20		0.29		0.23		0.37	
Observations	16,288		13,563		16,390		7458	
<i>Model 2: Mother's age at birth as a vector of dummy variables</i>								
Child's birth date	-0.10	0.09	-0.22	0.08	0.02	0.08	-0.12	0.12
Child's age-appropriate grade	0.01	0.04	-1.02	0.05***	0.25	0.04***	0.47	0.07***
Mother's AFQT score	0.26	0.02***	0.25	0.02***	0.26	0.02***	0.34	0.03***
Mother's years of education at birth	0.43	0.13***	0.24	0.13	0.52	0.12***	0.80	0.19***
Median family income, logged	1.96	0.36***	1.30	0.35***	1.59	0.33***	0.67	0.51
Mother unmarried at birth (1=yes)	0.10	0.74	-0.85	0.70	0.34	0.65	-1.21	0.96
Mother's age at birth = 16–19	1.78	2.31	1.23	4.64	-1.04	3.92	-0.32	3.58
Mother's age at birth = 20–23	2.08	2.30	1.92	4.63	-0.54	3.92	0.60	3.57
Mother's age at birth = 24–27	2.18	2.39	2.32	4.68	-0.03	3.97	0.39	3.69
Mother's age at birth = 28–31	3.16	2.53	4.14	4.74	0.66	4.03	1.02	3.88
Mother's age at birth = 32–36	3.71	2.70	3.60	4.80	-0.38	4.12	-1.18	4.13
Mother's age at birth = 36–39	6.62	3.10	9.03	5.04	1.15	4.35	-2.06	5.67
Race (black=1)	-12.09	13.33	-5.03	12.61	9.56	11.89	-8.65	17.14
Child's birth date × race	0.09	0.15	-0.03	0.13	-0.12	0.13	-0.27	0.20
Child's age-appropriate grade × race	-0.91	0.07***	-1.02	0.07***	-0.35	0.07***	0.26	0.11*
Mother's AFQT score × race	0.07	0.04*	0.08	0.03*	0.04	0.03	0.16	0.05**
Mother's years of education at birth × race	0.66	0.24**	0.55	0.21*	0.28	0.22	0.92	0.32**
Median family income × race	-0.89	0.55	-0.40	0.51	-0.87	0.49	0.22	0.77
Mother unmarried at birth (1=yes) × race	-0.57	0.96	0.26	0.87	-1.12	0.84	1.35	1.25
Mother's age at birth = 16–19 × race	-1.48	5.88	-0.58	6.71	-3.59	5.33	-5.37	5.33
Mother's age at birth = 20–23 × race	-4.33	5.89	-2.72	6.71	-5.01	5.35	-7.58	5.38
Mother's age at birth = 24–27 × race	-5.26	6.00	-3.43	6.79	-5.82	5.45	-8.17	5.59
Mother's age at birth = 28–31 × race	-4.53	6.16	-2.90	6.89	-3.70	5.58	-3.04	5.89
Mother's age at birth = 32–36 × race	-3.97	6.39	-1.08	7.05	-2.85	5.78	-2.11	6.45
Mother's age at birth = 36–39 × race	-4.25	7.36	-2.78	7.89	-6.17	6.62	1.51	8.15
Constant	73.57	7.63	91.87	8.24	63.89	7.58	62.08	10.79
<i>R</i> ² within	0.04		0.19		0.00		0.03	
<i>R</i> ² between	0.26		0.33		0.31		0.39	
<i>R</i> ² overall	0.21		0.29		0.23		0.37	
Observations	16,288		13,563		16,390		7458	

(continued on next page)

Table 3 (continued)

Variables by model	PIAT-RR		PIAT-RC		PIAT-M		PPVT-R	
	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.
<i>Model 3a: Black and white observations matched on mother's age at birth, birth age as a continuous variable</i>								
Child's birth date	-0.14	0.13	-0.36	0.13**	0.00	0.13	-0.09	0.15
Child's age-appropriate grade	-0.31	0.09***	-1.39	0.11***	0.20	0.09*	0.38	0.09***
Mother's AFQT score	0.26	0.03***	0.23	0.03***	0.26	0.03***	0.34	0.03***
Mother's years of education at birth	0.37	0.19	0.30	0.20	0.41	0.18*	0.53	0.22*
Median family income, logged	2.73	0.49***	2.04	0.46***	2.00	0.46***	1.23	0.56*
Mother unmarried at birth (1=yes)	0.62	0.92	-0.60	0.89	0.46	0.81	-1.65	1.00
Mother's age at birth	0.20	0.15	0.20	0.14	0.21	0.14	-0.02	0.16
Race (black=1)	-8.96	13.54	1.64	13.36	10.04	12.68	-24.58	15.41
Child's birth date × race	0.04	0.20	-0.16	0.19	-0.19	0.19	-0.22	0.23
Child's age-appropriate grade × race	-0.90	0.13***	-1.12	0.15***	-0.12	0.13	0.23	0.12
Mother's AFQT score × race	0.10	0.05*	0.12	0.04**	0.02	0.04	0.17	0.05***
Mother's years of education at birth × race	0.56	0.30	0.40	0.29	0.42	0.28	1.07	0.34**
Median family income × race	-1.38	0.71	-1.09	0.66	-1.26	0.64*	-0.37	0.80
Mother unmarried at birth (1=yes) × race	-0.40	1.16	1.24	1.10	-1.44	1.04	1.91	1.29
Mother's age at birth × race	-0.11	0.22	0.09	0.20	0.05	0.20	0.16	0.24
Constant	74.56	8.81	101.60	9.15	60.07	8.31	62.93	9.76
R^2 within		0.09		0.27		0.00		0.03
R^2 between		0.23		0.27		0.26		0.38
R^2 overall		0.21		0.27		0.21		0.36
Observations		5628		4748		5669		5832
<i>Model 3b: Black and white observations matched on mother's age at birth, birth age as a vector of dummy variables</i>								
Child's birth date	-0.15	0.12	-0.33	0.12	0.06	0.11	-0.11	0.14
Child's age-appropriate grade	-0.29	0.09***	-1.38	0.11***	0.18	0.09*	0.35	0.09***
Mother's AFQT score	0.26	0.03***	0.24	0.03***	0.26	0.03***	0.34	0.03***
Mother's years of education at birth	0.39	0.19*	0.32	0.20	0.43	0.18*	0.44	0.22*
Median family income, logged	2.71	0.49***	2.05	0.46***	2.05	0.46***	1.25	0.56*
Mother unmarried at birth (1=yes)	0.58	0.93	-0.57	0.90	0.42	0.81	-1.48	1.01
Mother's age at birth=16–19	8.22	3.46	5.71	3.56	-1.52	4.72	-0.24	3.72
Mother's age at birth=20–23	8.52	3.47	6.47	3.57	-0.87	4.72	1.05	3.73
Mother's age at birth=24–27	8.14	3.58	6.57	3.71	-0.40	4.80	1.15	3.87
Mother's age at birth=28–31	9.96	3.77	7.65	3.88	0.78	4.91	2.25	4.14*
Mother's age at birth=32–36	12.85	4.15	8.82	4.46	-0.80	5.29	-1.33	4.62**
Mother's age at birth=36–39	16.37	5.14	22.83	4.94	0.64	6.66	-6.07	7.12***
Race (black=1)	-1.21	16.49	10.00	15.53	12.40	16.64	-12.89	17.66
Child's birth date × race	0.09	0.19	-0.11	0.18	-0.24	0.17	-0.29	0.21
Child's age-appropriate grade × race	-0.91	0.13	-1.13	0.15	-0.11	0.13	0.26	0.13
Mother's AFQT score × race	0.09	0.05	0.11	0.04	0.01	0.04	0.16	0.05
Mother's years of education at birth × race	0.71	0.30	0.56	0.28	0.47	0.28	1.28	0.34
Median family income × race	-1.41	0.71	-1.11	0.66	-1.29	0.64	-0.47	0.80
Mother unmarried at birth (1=yes) × race	-0.85	1.17	0.76	1.11	-1.57	1.04	1.49	1.29
Mother's age at birth=16–19 × race	-13	7	-9	6	3	9	-1	5
Mother's age at birth=20–23 × race	-15.85	7.15	-11.26	5.72	2.13	9.31	-4.12	4.80
Mother's age at birth=24–27 × race	-16.53	7.28	-11.69	5.88	1.47	9.40	-5.17	5.05
Mother's age at birth=28–31 × race	-15.71	7.47	-9.49	6.09	4.20	9.52	-0.49	5.45
Mother's age at birth=32–36 × race	-16.20	7.96	-4.09	6.73	5.38	9.89	1.56	6.22
Mother's age at birth=36–39 × race	-19.23	9.89	2.97	7.05	2.37	11.60	8.25	8.86
Constant	72.02	10.39	97.01	10.53	60.35	10.32	63.70	11.81
R^2 within		0.09		0.27		0.00		0.03
R^2 between		0.24		0.28		0.26		0.39
R^2 overall		0.22		0.27		0.22		0.37
Observations		5628		4748		5669		5832

*** denotes $p < .001$, ** denotes $p < .01$, * denotes $p < .05$.

Table 4
Net change in the B–W difference over time

Model	PIAT-RR	PIAT-RC	PIAT-M	PPVT-R
<i>b for the race × birth date interaction term</i>				
1	0.00	–0.13	–0.14	–0.20
2	0.09	–0.03	–0.12	–0.27
3a	0.04	–0.16	–0.19	–0.22
3b	0.09	–0.11	–0.24	–0.29
<i>Implied change in the B–W difference per decade in S.D.s (S.D. = 15)</i>				
1	0.00	0.09	0.09	0.13
2	–0.06	0.02	0.08	0.18
3a	–0.03	0.11	0.13	0.15
3b	–0.06	0.07	0.16	0.19

None of the coefficients is significant at the .05 level.

Variables: Child's birth date, child's age-appropriate grade, and family income; mother's IQ, education, marital status, and age at birth; dummy variable for race; interaction terms between race and the other independent variables.

terms of the implied change in the B–W difference per decade, expressed in standard deviations.

The changes are small and statistically consistent with an interpretation of “no change” in the B–W difference, but the sign of the coefficient for the interaction term is consistent within tests. All four specifications of the reading recognition test indicated a small convergence in the B–W difference (even the coefficient that rounded to .00 in Model 1 was positive at the third decimal place). All four specifications of the other three tests indicated a small increase in the B–W difference. The largest increase, still not reaching statistical significance, was for the PPVT-R, where the implied increases in the B–W difference ranged from .13 to .19 S.D.s per decade in the four versions of the analysis.

It is noteworthy that the dummy variable for race was not statistically significant in any of the models, indicating the degree to which the race interaction terms soaked up the sources of the raw B–W difference. However, this result cannot be interpreted as meaning that race becomes insignificant after controlling for child's birth date, age, and the family background variables. If the interaction terms are omitted, the coefficient for a dummy variable for race is significant beyond the .001 level for all of the tests but reading recognition in all specifications of the model.

The results in Table 4 supplement previous evidence from the NAEP and SAT that no closure in the B–W difference occurred among children born from the early 1970s into the last half of the 1990s on achievement tests. The results are inconsistent with the narrowing of the B–W difference in the IQ test standardizations during the same period found by Dickens and Flynn (in press-a). This is most explicitly true for the PPVT-R, a measure of

verbal IQ, but the failure of the PIAT-RC and PIAT-M to converge is also relevant, insofar as both tests measure reasoning ability along with learned knowledge.

Potential technical explanations for this inconsistency were explored. One possibility was that the selection of NLSY children for testing created an artifact. The NLSY tried to test all children who were eligible (because of their age or prior testing history) for a given test on a given test year, but, after the initial 1986 test wave, it was decided to test only children who were living full-time or part-time with their mothers. This criterion meant that proportionately more black than white children were ineligible for testing (15.2% of testing opportunities for black children compared to 10.2% for white children). Not living with the mother is associated with factors that might tend to depress test scores (abandonment, foster care, institutionalization). The potentially lowest-scoring black children were in this respect probably under-represented in the NLSY sample. But this artifact would tend to understate the real B–W test differences for the test waves following 1986, because the omitted children, disproportionately disadvantaged, were also disproportionately black. No other inconsistencies in the selection of black and white children for testing were identified.

A second possibility was that the sample for this analysis was systematically skewed by including cases that were part of the NLSY's oversampling of blacks and low-income whites. To explore this possibility, Models 1 and 2 were replicated with observations restricted to the 13,602 observations of children of NLSY women who were part of the nationally representative cross-sectional sample. The results are shown in Appendix Table A-2, but they may be summarized quickly: None of the differences in the results from the two samples approached statistical significance. The results leave open the possibility that profiles of the B–W difference vary by socioeconomic class, but that topic requires a full-scale analysis of its own.

A third possibility was that the use of birth date as the independent variable of interest affected the results. Age at testing had a significant interaction effect with race, which could overlap with the effect of the interaction between race and birth date, raising potential multicollinearity problems. Accordingly, all of the analyses were replicated substituting test year for birth date in the GLS regressions. The results for Models 1 and 2 are also shown in Table A-2. Without exception, the interaction between race and test year produces coefficients that are within a few hundredths of the coefficients for the interaction between race and birth date.

A fourth possibility was that all of the models control for too much. If the socioeconomic position of black NLSY mothers improved relative to white mothers during

the observation period, controlling for the family background variables could mask environmentally-caused improvements in the performance of the children. Replications of the interaction analyses using reduced models that omitted mother's AFQT score, education at birth, marital status at birth, and family income were conducted to test this possibility. The results from Models 1 and 2 are shown in Table A-2. All of the negative coefficients are larger when the family background variables are omitted, and the sign of the coefficient for reading recognition changes from positive to negative.

Finally, it may be asked whether the dummy variables used for mother's age at birth, consisting of 4-year segments, were sufficiently narrow to preclude multicollinearity problems within the segments. Three versions of Model 2 were explored, dividing the mother's age at birth into 2-year and 1-year segments as well as 4-year segments. Those results are also summarized in Table A-2. The version with the 4-year segments presented in the text produced the smallest estimates of increases in the B–W difference.

4. Discussion

We are faced with a familiar story of inconsistent findings. Huang and Hauser (2001) and Dickens and Flynn (in press-a) present plausible evidence that the B–W difference on cognitive tests has narrowed, while the present article presents plausible evidence that it did not narrow at all and possibly widened.

The results presented here cannot resolve the issue, but they must be given some weight in assessing it. The black sample sizes for the NLSY children are very large compared to those used in IQ standardization samples, which typically consist of just a few hundred children (e.g., the black sample for the 2002 standardization of the WISC-IV consisted of 343 subjects [Prifitera, Weiss, Saklofske, & Rolfhus, 2005]). Most of the subjects were tested more than once. The testing procedures identify and test children who are easily missed in surveys that rely on children being present at school on a day designated for testing. The same versions of the tests were used across time. Cognitive test scores were available for the mothers, along with extensive family background data. While no single database can provide dispositive results, the database for the NLSY79 children has many strengths, and it yields consistent evidence, robust across alternative specifications and samples, that the B–W difference did not diminish on either academic achievement or cognitive tests for children born from the mid 1970s through the mid 1990s.

A useful next step in this dialogue would be to explore the possibility that we are looking at a larger picture with the occasional anomaly but an identifiable overall pattern.

This hypothesis starts with assumption that an important distinction between achievement tests and IQ tests applies only to the most focused tests of learned material. More general tests such as the NAEP, SAT, and the PIAT battery are sufficiently *g* loaded that their trends over time can be aggregated with the results from IQ tests. If so, then the evidence from the NAEP and SAT in the 1970s and 1980s points to a genuine reduction in the B–W IQ difference, and the subsequent plateau constitutes genuine evidence that narrowing in the IQ gap has stopped. Any estimate of the size of the reduction must be hedged with qualifications, but Dickens' and Flynn's (in press-a) estimate of 3–6 IQ points from a base of about 16–18 points is a useful, though provisional, starting point.

The results presented here reinforce the evidence for the plateau, but they do not contradict the evidence for the earlier narrowing. The findings in the present article do not contradict the Huang and Hauser (2001) birth cohort analysis of the General Social Survey. Ninety-seven percent of the GSS subjects who produced their results were born before the first NLSY child was born. Nor do the findings in the present article contradict the reduction in the B–W difference found in the NAEP and SAT. All of the narrowing in the NAEP difference occurred among cohorts born no later than 1979, with the great majority of the children in those cohorts born prior to 1975. All of the narrowing in the SAT difference occurred among cohorts born prior to 1975. The NLSY children in the present study were born during exactly the same period, from the mid 1970s onward, as the children who produced the plateaus in the NAEP since the 1980s and in the SAT since 1991.

The inconsistency between the evidence from the NLSY and from the IQ standardizations remains, but it is not necessarily irreconcilable. The Dickens and Flynn (in press-a) evidence for a narrowing gap was based on means and standard deviations for age groups. The hypothesis suggested by the data from the NLSY children is that if individual data from the IQ standardizations were analyzed by birth cohort, they would reveal that the effect that Dickens and Flynn found was concentrated among subjects born before the late 1970s. Such a finding would inform a variety of theoretical explanations of the B–W difference, but elaborating on those implications is premature until the hypothesis is tested.

Acknowledgments

The author wishes to thank William Dickens, Greg Duncan, and Douglas Hibbs for helpful commentary on early drafts of this manuscript, and three reviewers for critiques that materially improved the submitted version.

Appendix A

Table A-1
Full results from the GLS analysis of white and black trends over time

Model and Variables	PIAT-RR				PIAT-RC				PIAT-M				PPVT-R			
	White		Black		White		Black		White		Black		White		Black	
	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.
<i>Model 1: Mother's age at birth as a continuous variable</i>																
Child's birth date	-0.06	0.10	-0.06	0.13	-0.20	0.09*	-0.33	0.11**	0.02	0.09	-0.12	0.11	-0.11	0.13	-0.31	0.18
Child's age-appropriate grade	0.01	0.04	-0.90	0.05***	-1.03	0.05***	-2.06	0.05***	0.26	0.04***	-0.10	0.05*	0.49	0.07***	0.72	0.08***
Mother's AFQT score	0.26	0.02***	0.35	0.03***	0.26	0.02***	0.35	0.03***	0.26	0.02***	0.30	0.03***	0.34	0.03***	0.51	0.04***
Mother's years of education at birth	0.42	0.13***	0.91	0.20***	0.22	0.13	0.62	0.17***	0.53	0.12***	0.68	0.18***	0.84	0.19***	1.55	0.26***
Median family income from birth to age 4, logged	1.98	0.36***	1.15	0.41**	1.32	0.35***	0.95	0.38*	1.61	0.33***	0.76	0.36*	0.68	0.51	0.98	0.57
Mother unmarried at birth (1=yes)	0.14	0.74	-0.05	0.60	-0.78	0.70	-0.21	0.52	0.33	0.64	-0.57	0.53	-1.34	0.94	0.41	0.81
Mother's age at birth	0.11	0.11	0.08	0.14	0.21	0.10	0.25	0.11	0.07	0.09	0.15	0.11	-0.06	0.14	0.17	0.17
Constant	70.4	6.45	63.0	8.57	87.7	6.12	89.1	7.51	61.4	5.84	66.4	7.63	62.5	8.90	37.2	11.83
<i>R</i> ² within	0.00		0.09		0.09		0.30		0.01		0.00		0.02		0.05	
<i>R</i> ² between	0.18		0.21		0.20		0.30		0.20		0.18		0.15		0.23	
<i>R</i> ² overall	0.13		0.18		0.17		0.29		0.14		0.12		0.14		0.22	
Observations	9573		6715		7913		5650		9633		6757		4287		3171	
<i>Model 2: Mother's age at birth as a vector of dummy variables</i>																
Child's birth date	-0.10	0.09	-0.01	0.12	-0.22	0.08**	-0.24	0.10*	0.02	0.08	-0.10	0.10	-0.12	0.12	-0.39	0.16*
Child's age-appropriate grade	0.01	0.04	-0.89	0.05***	-1.02	0.05***	-2.05	0.05***	0.25	0.04***	-0.10	0.05	0.47	0.07***	0.73	0.08***
Mother's AFQT score	0.26	0.02***	0.34	0.03***	0.25	0.02***	0.33	0.03***	0.26	0.02***	0.29	0.03***	0.34	0.03***	0.50	0.04***
Mother's years of education at birth	0.43	0.13***	1.08	0.20***	0.24	0.13	0.78	0.17***	0.52	0.12***	0.80	0.18***	0.80	0.19***	1.72	0.26***
Median family income, logged	1.96	0.36***	1.06	0.41***	1.30	0.35***	0.90	0.37*	1.59	0.33***	0.73	0.36*	0.67	0.51	0.88	0.57
Mother unmarried at birth (1=yes)	0.10	0.74	-0.47	0.60	-0.85	0.70	-0.59	0.52	0.34	0.65	-0.78	0.53	-1.21	0.95	0.13	0.80
Mother's age at birth=16-19	1.77	2.31	0.31	5.41	1.23	4.65	0.65	4.85	-1.03	3.92	-4.62	3.60	-0.32	3.58	-5.68	3.96
Mother's age at birth=20-23	2.07	2.29	-2.25	5.43	1.92	4.64	-0.80	4.86	-0.54	3.91	-5.54	3.65	0.60	3.57	-6.97	4.03
Mother's age at birth=24-27	2.17	2.39	-3.08	5.50	2.32	4.68	-1.10	4.92	-0.03	3.97	-5.84	3.73	0.40	3.69	-7.77	4.20
Mother's age at birth=28-31	3.15	2.53	-1.36	5.62	4.14	4.74	1.24	5.01	0.66	4.03	-3.04	3.85	1.02	3.88	-2.03	4.44
Mother's age at birth=32-36	3.71	2.70	-0.25	5.79	3.60	4.81	2.53	5.16	-0.38	4.12	-3.23	4.05	-1.18	4.13	-3.29	4.96
Mother's age at birth=36-39	6.62	3.10*	2.39	6.68	9.03	5.04	6.24	6.07	1.16	4.35	-5.02	4.99	-2.06	5.67	-0.55	5.85
Constant	73.6	7.63	61.4	10.93	91.8	8.24	86.9	9.54	63.9	7.58	73.4	9.16	62.1	10.79	53.4	13.32
<i>R</i> ² within	0.00		0.09		0.09		0.30		0.01		0.00		0.02		0.05	
<i>R</i> ² between	0.18		0.22		0.20		0.31		0.20		0.19		0.15		0.25	
<i>R</i> ² overall	0.13		0.18		0.17		0.30		0.14		0.13		0.14		0.23	
Observations	9573		6715		7913		5650		9633		6757		4287		3171	

(continued on next page)

Table A-1 (continued)

Model and Variables	PIAT-RR				PIAT-RC				PIAT-M				PPVT-R			
	White		Black		White		Black		White		Black		White		Black	
	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.	<i>b</i>	Robust S.E.
<i>Model 3a: Black and white observations matched on mother's age at birth, birth age as a continuous variable</i>																
Child's birth date	-0.14	0.13	-0.10	0.15	-0.36	0.13**	-0.51	0.14***	0.00	0.13	-0.19	0.14	-0.09	0.15	-0.31	0.18
Child's age-appropriate grade	-0.31	0.09***	-1.21	0.09***	-1.39	0.11***	-2.51	0.10***	0.20	0.09*	0.08	0.09	0.38	0.09***	0.61	0.09***
Mother's AFQT score	0.26	0.03***	0.36	0.04***	0.23	0.03***	0.36	0.03***	0.26	0.03***	0.28	0.03***	0.34	0.03***	0.51	0.04***
Mother's years of education at birth	0.37	0.19	0.93	0.23***	0.30	0.20	0.70	0.21***	0.41	0.18*	0.82	0.21***	0.53	0.22*	1.60	0.27***
Median family income, logged	2.73	0.49***	1.34	0.52**	2.04	0.46***	0.95	0.47*	2.00	0.46***	0.74	0.45	1.22	0.56*	0.86	0.57
Mother unmarried at birth (1=yes)	0.63	0.92	0.22	0.71	-0.61	0.89	0.63	0.65	0.46	0.81	-0.98	0.65	-1.66	1.01	0.26	0.81
Mother's age at birth	0.20	0.15	0.09	0.16	0.20	0.14	0.29	0.14	0.21	0.14	0.26	0.14	-0.02	0.16	0.14	0.18
Constant	74.6	8.81	65.6	10.29	101.7	9.16	103.2	9.73	60.1	8.32	70.1	9.58	63.1	9.77	38.4	11.92
R^2 within	0.01		0.14		0.16		0.33		0.00		0.00		0.01		0.04	
R^2 between	0.15		0.21		0.14		0.26		0.17		0.15		0.14		0.23	
R^2 overall	0.14		0.19		0.15		0.27		0.15		0.12		0.13		0.21	
Observations	2796		2832		2343		2405		2814		2855		2916		2916	
<i>Model 3b: Black and white observations matched on mother's age at birth, birth age as a vector of dummy variables</i>																
Child's birth date	-0.16	0.12	-0.07	0.14	-0.33	0.12**	-0.44	0.13***	0.06	0.11	-0.18	0.13	-0.11	0.14	-0.40	0.16*
Child's age-appropriate grade	-0.29	0.09*	-1.20	0.09***	-1.38	0.11***	-2.52	0.10***	0.18	0.09*	0.07	0.09	0.35	0.09***	0.61	0.09***
Mother's AFQT score	0.26	0.03***	0.35	0.04***	0.24	0.03***	0.35	0.03***	0.26	0.03***	0.28	0.03***	0.34	0.03***	0.50	0.04***
Mother's years of education at birth	0.39	0.19*	1.11	0.23***	0.32	0.20	0.88	0.20***	0.43	0.18*	0.90	0.21***	0.44	0.22*	1.72	0.26***
Median family income, logged	2.71	0.49***	1.30	0.51*	2.05	0.46***	0.94	0.47*	2.05	0.46***	0.75	0.45	1.25	0.56	0.79	0.57
Mother unmarried at birth (1=yes)	0.59	0.93	-0.27	0.71	-0.57	0.90	0.20	0.65	0.42	0.81	-1.14	0.65	-1.49	1.02	0.01	0.80
Mother's age at birth=16–19	8.22	3.46*	-4.46	6.25	5.71	3.56	-3.37	4.45	-1.52	4.72	1.58	8.02	-0.23	3.72	-1.45	2.96
Mother's age at birth=20–23	8.52	3.47*	-7.28	6.26	6.48	3.57	-4.83	4.47	-0.87	4.72	1.30	8.04	1.06	3.73	-3.08	3.02
Mother's age at birth=24–27	8.14	3.58*	-8.34	6.34	6.58	3.71	-5.15	4.56	-0.41	4.80	1.10	8.09	1.15	3.87	-4.02	3.24
Mother's age at birth=28–31	9.97	3.77**	-5.68	6.45	7.66	3.88*	-1.88	4.69	0.78	4.91	5.02	8.16	2.27	4.14	1.75	3.54
Mother's age at birth=32–36	12.86	4.15**	-3.30	6.79	8.84	4.46*	4.69	5.03	-0.81	5.29	4.61	8.36	-1.31	4.62	0.23	4.16
Mother's age at birth=36–39	16.38	5.14***	-2.80	8.45	22.85	4.94*	25.75	5.02*	0.64	6.66	3.05	9.51	-6.06	7.12	2.18	5.27
Constant	72.0	10.39	70.7	12.81	97.1	10.55	107.1	11.40	60.4	10.32	72.7	13.06	63.8	11.82	50.9	13.13
R^2 within	0.01		0.14		0.16		0.33		0.00		0.00		0.01		0.04	
R^2 between	0.16		0.22		0.15		0.28		0.17		0.15		0.14		0.24	
R^2 overall	0.14		0.20		0.15		0.28		0.15		0.13		0.14		0.23	
Observations	2796		2832		2343		2405		2814		2855		2916		2916	

*** denotes $p < .001$, ** denotes $p < .01$, * denotes $p < .05$.

Table A-2

The net change in the B–W difference over time using alternative samples and independent variables

I. Comparison of results from the full sample and the children of women in the cross-sectional sample

	<i>b</i> for the race×birth date interaction term			
	PIAT-RR	PIAT-RC	PIAT-M	PPVT-R
Model 1				
Full sample	0.00	−0.13	−0.14	−0.20
Cross-sectional sample	0.06	−0.06	−0.21	−0.07
Model 2				
Full sample	0.09	−0.03	−0.12	−0.27
Cross-sectional sample	0.20	0.13	−0.15	−0.15

II. Comparison of results when test year is substituted for child's birth date using the full sample

	<i>b</i> for the interaction term			
	PIAT-RR	PIAT-RC	PIAT-M	PPVT-R
Model 1				
Race×birth date	0.00	−0.13	−0.14	−0.20
Race×test year	−0.02	−0.18	−0.17	−0.22
Model 2				
Race×birth date	0.09	−0.03	−0.12	−0.27
Race×test year	0.07	−0.07	−0.15	−0.29

III. Comparison of results when the family background variables are omitted

	<i>b</i> for the race×birth date interaction term			
	PIAT-RR	PIAT-RC	PIAT-M	PPVT-R
Model 1				
Using all covariates	0.00	−0.13	−0.14	−0.20
Using just birth date, age, mother's age at birth, and race	−0.23	−0.33*	−0.25*	−0.33
Model 2				
Using all covariates	0.09	−0.03	−0.12	−0.27
Using just birth date, age, mother's age at birth, and race	−0.11	−0.20	−0.21	−0.32

IV. Comparison of results from four specifications of mother's age at birth in Model 2

	<i>b</i> for the race×birth date interaction term			
	PIAT-RR	PIAT-RC	PIAT-M	PPVT-R
Mothers age at birth treated as a continuous variable, full sample	0.00	−0.13	−0.14	−0.20
Mother's age at birth grouped into 4-year segments, using full sample	0.09	−0.03	−0.12	−0.27
Mother's age at birth grouped into 2-year segments, using full sample	−0.14	−0.22	−0.35	−0.34
Mother's age at birth grouped into 1-year segments, for children born to mothers ages 15 through 34	−0.08	−0.20	−0.20	−0.30

*** denotes $p < .001$, ** denotes $p < .01$, * denotes $p < .05$.

References

- Baker, P. C., Keck, C. K., Mott, F. L., & Quinlan, S. V. (1993). *NLSY child handbook, rev. edition: A guide to the 1986–1990 national longitudinal survey of youth child data*. Columbus, Ohio: Center for Human Resource Research, The Ohio State University.
- College Board. (annual). *Profiles, college bound seniors*. New York: College Entrance Examination Board.
- Dickens, W. T., & Flynn, J. R. (in press-a). Black Americans reduce the racial IQ gap: Evidence from standardization samples. *Psychological Science*.
- Dickens, W. T., & Flynn, J. R. (in press-b). Common ground and differences. *Psychological Science*.
- Dunn, L. M., & Markwardt, F. C. (1970). *Peabody individual achievement test*. Circle Pines, MN: American Guidance Service.
- 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*. New York: Wiley.
- Grutter v. Bollinger et al., 539 U.S. 306 (2003).
- Hedges, L. V., & Nowell, A. (1998). Black–white test score convergence since 1965. In C. Jencks, & M. Phillips (Eds.), *The black–white test score gap* (pp. 149–181). Washington: Brookings Institution Press.
- Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.
- Huang, M. -H., & Hauser, R. M. (2001). Convergent trends in black–white verbal test score differentials in the U.S.: Period and cohort perspectives. *EurAmerica*, 31(2), 185–230.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Lynn, R. (1998). Has the black–white intelligence difference in the United States been narrowing over time? *Personality and Individual Differences*, 25, 999–1002.
- Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence*. New York: Oxford University Press.
- Perie, M., & Moran, R. (2005). *NAEP 2004 trends in academic progress: Three decades of student performance in reading and mathematics*. Washington: National Center for Education Statistics.
- Prifitera, A., Weiss, L. G., Saklofske, D. H., & Rolfhus, E. (2005). The WISC-IV in the clinical assessment context. In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist–practitioner perspectives* (pp. 3–32).
- Robertson, G. J., & Eisenberg, J. L. (1981). *Technical supplement, forms L and M, Peabody picture vocabulary test-revised*. Circle Pines, MN: American Guidance Service.
- Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., III, & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, 54, 297–330.
- Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, 11(2), 235–294.
- Rushton, J. P., & Jensen, A. R. (in press). The totality of available evidence shows race-IQ gap still remains. *Psychological Science*.
- Vincent, K. R. (1991). Black/white IQ differences: Does age make the difference? *Journal of Clinical Psychology*, 47(2), 266–270.