Convergent Trends in Black-White Test-Score Differentials in the U.S.: A Correction of Richard Lynn[1]

Min-Hsiung Huang
Institute of European and American Studies
Academia Sinica, Taipei, Taiwan

Robert M. Hauser
Center for Demography and Ecology
Department of Sociology
The University of Wisconsin-Madison

Rev. July 31, 2000

[1] Support for this research was provided by the Institute of European and American Studies, Academia Sinica, Taipei, Taiwan, and by the Vilas Estate Trust and the Center for Demography and Ecology at the University of Wisconsin-Madison. Data from the General Social Survey are available from the Interuniversity Consortium for Political and Social Research, University of Michigan, Ann Arbor, Michigan, or at http://www.icpsr.umich.edu/gss/. The opinions expressed herein are those of the authors, and the data files used in this analysis are available directly from the authors. Address correspondence to Min-Hsiung Huang, Institute of European and American Studies, Academia Sinica, Taipei 11529, Taiwan, or E-MAIL to MHHUANG@SINICA.EDU.TW.

Abstract

Using aggregate data from the General Social Survey (GSS), 1974-1996, Lynn (1998) claims that the Black-White intelligence difference in the United States has not been narrowing over time. We replicate Lynn's analysis and challenge his conclusion by identifying several methodological problems. By analyzing changes in Black-White differences in the GSS vocabulary test across survey years, rather than birth cohorts, Lynn overlooks both the duration and the significance of the Black-White convergence. We extend our earlier intercohort analysis of GSS data through 1998 and confirm our previous finding of a very significant, long-term Black-White convergence, which is attributable to improvements in socioeconomic background and schooling among African-Americans. Even in an analysis of aggregate changes in the Black-White test score gap across survey years, when data are weighted properly to represent the U.S. population on the individual level, we find that the Black-White test score gap narrowed significantly over the period from 1974 to 1998.

INTRODUCTION

Using data from the General Social Survey (GSS), 1974-1994, Huang and Hauser (1998) conclude that Black-White differences in GSS vocabulary test scores (WORDSUM) have been decreasing over time. However, using aggregate WORDSUM data from the GSS, 1974-1996, Lynn (1998) maintains that the Black-White intelligence difference in the US has not been narrowing over time.[1] To reconcile the inconsistencies in findings, we attempted to replicate Lynn's (1998) analysis. We identified several methodological problems in Lynn's study, and we extended Huang and Hauser's (1998) inter-cohort analysis of Black-White differences in WORDSUM for two more rounds of the survey, that is, using data from the GSS from 1974 to 1998.[2]

REPLICATING LYNN'S ANALYSIS

The GSS is a series of repeated cross-sectional surveys of the adult population in households in the United States. A ten-item multiple-choice vocabulary test was administered to respondents in the GSS survey years of 1974, 1976, 1978, 1982, 1984, 1987, 1988, 1989, 1990, 1991, 1993, 1994, 1996, and 1998.[3] The variable, WORDSUM, ranges from 0 to 10, denoting the total number of test items answered correctly. The data for Lynn's (1998) study come from the GSS, 1974-1996. Lynn reports there are 14,657 Whites and 2,510 Blacks. While we have reproduced his count for Blacks,

---

[1] Lynn (1998) regards the GSS vocabulary test as a measure of intelligence.

[2] A draft of our 1998 paper was presented at an Emory University conference in April 1996, which Lynn attended (Hauser and Huang 1996, Huang and Hauser 1996). Curiously, Lynn's 1998 paper does not refer either to our work with the General Social Survey or to the conference volume (Neisser 1998).

[3] See Huang and Hauser (1998:304-11) and Hauser and Huang (1997:340-45) for a description and analyses of the properties of WORDSUM.

we can not reproduce his exact count for Whites.[4] Starting with the full file and excluding individuals to whom WORDSUM was not administered, we count 15,281 Whites and 2,807 Blacks, for a total of 18,088. There are several ways of proceeding from that figure, none of which gives us Lynn's numbers for both Whites and Blacks.

First, we deleted those cases in which WORDSUM=99. This gave us 14,613 Whites and exactly 2,510 Blacks. This method gives us Lynn's count of Blacks, but leaves out 44 Whites. Our total adds up to 17,123, while Lynn computes a total of 17,167. Second, we deleted those cases in which the responses are "9" (no answer) for all 10 word items, recoding 9 as 0 for each word item, and reconstructing WORDSUM by adding up the correct item responses. This gave us 14,633 Whites and 2,504 Blacks, for a total of 17,137. Finally, by following the second procedure, we found that several of the deleted cases have WORDSUM = 0, rather than 99. If we add these cases back in, we get 14,649 Whites and 2,516 Blacks, for a total of 17,165 cases. None of these three case selection procedures yields the exact count for Whites that is recorded in Lynn's sample. To repeat Lynn's analysis, we chose to use the first case selection procedure, which yields 14,613 Whites and exactly 2,510 Blacks.

Lynn (1998:1001) uses "*d* scores" to express Black-White differences in vocabulary scores. The year-specific *d* scores denote the differences between the mean scores of Whites and the mean scores of Blacks, divided by the "pooled standard deviation." Lynn (1998) does not mention whether

---

[4] We contacted Richard Lynn by letter to learn how he had selected GSS cases, and in March 1999, he wrote to say that he had discarded his work with the WORDSUM data and could not respond to our queries.

or not he uses the pooled across-year standard deviation or the pooled within-year standard deviation. Therefore, Table 1 displays five versions of *d* value by survey year. The *d* values in the first column are as reported by Lynn (1998). Unweighted *d* values based on the pooled across-year standard deviation are presented in column 2, while unweighted *d* values based on the pooled within-year standard deviation are presented in column 4. In order for the GSS samples to be representative of the U.S. population on the individual level, and to take into account the over-sampling of blacks in the 1982 and 1987 surveys, it is necessary to weight the data by the product of two variables in the public GSS data file, ADULTS (the number of adults in a household) and OVERSAMP (Davis and Smith, 1992). The *d* values based on weighted means and weighted pooled across-year standard deviations are presented in column 3, while column 5 presents the *d* values based on weighted means and weighted pooled within-year standard deviations. The *d* values in column 2 are closest to Lynn's *d* values; the figures agree to two decimal places in six years.[5] This suggests that Lynn may have incorrectly used the unweighted data to estimate pooled across-year standard deviations.

According to Lynn (1998), the Pearson correlation between GSS survey year and the size of Black-White difference (*d* value) is r = -.32 (p = 0.28), which implies a narrowing gap but is not statistically significant.[6] As shown in the last row of the second column of Table 1, our replication of Lynn's analysis, based on the unweighted estimates of *d* in the second column, yields a higher correlation of r = -.41 (p =0.17), but the correlation is still not statistically significant. When data are

---

[5] Note that none of our four estimates is at all close to Lynn's *d* = 0.56 in 1993.

[6] Note that Lynn's *p*-values are perhaps excessively conservative, based on a two-tailed test. If the alternative hypothesis were convergence, one could properly choose a one-tailed test, yielding a *p*-value half as large as reported by Lynn.

weighted, however, as shown in the last row of the third column of Table 1, the Pearson correlation between the GSS survey year and the size of Black-White difference (*d* value) reaches -0.55 (p = 0.052), reaching statistical significance at the .03 level in a one-tailed test, but not quite at p = .05 in a two-tailed test. Modelled on Table 1, Table 2 displays the *d* values when data from the 1998 GSS are added. As indicated in the last row of the third column of Table 2, when data are weighted and when data from the 1998 GSS are added, a significant correlation of r = -0.64 (p = 0.01) suggests that the Black-White difference in WORDSUM narrowed over the 24 year period from 1974 to 1998. On the other hand, when *d* values are based on the pooled within-year standard deviation, with or without weighting, the magnitude of the narrowing Black-White gap in WORDSUM becomes insignificant, as indicated in the last rows of columns 4 and 5 of Table 1 and Table 2. This is because the pooled within-year standard deviation of WORDSUM decreases over time. It decreases from 2.2 in 1974 to 2.0 in 1998, as persons who score extremely high or extremely low become a proportionally smaller share of the sample.[7] Consistent with Lynn's analysis, we think that the *d* values based on the pooled across-year standard deviation should be used.

Lynn (1998:1001) also regresses year of survey on *d* values, and we have done the same. The regression results for the series in Table 1 and Table 2 are presented in Table 3. When weighting is applied and when *d* values are based on the pooled across-year standard deviation, year of survey is

---

[7] An increase for Blacks in test scores and a decrease in the proportion of Whites who score extremely high may contribute to the finding that the pooled within-year standard deviation of WORDSUM tends to decrease over time. As the standard deviation of WORDSUM decreases, the *d* values based on the pooled within-year standard deviations tend to increase. Therefore, the *d* values based on the pooled within-year standard deviations tend to conceal Black-White convergence in test scores across time.

found to have significant negative effects on the size of Black-White difference in WORDSUM. Lynn (1998) does not find a narrowing gap across the 22-year period between 1974 and 1996, partly because the data are not weighted. But why does the narrowing gap in the size of Black-White difference in WORDSUM become more significant when data are weighted? We will address this question later.

BLACK-WHITE DIFFERENCE IN WORDSUM ACROSS SURVEY YEAR

A more rigorous way to test whether there has been a statistically significant narrowing gap between Blacks and Whites in vocabulary scores across survey year is not to use *d* scores, but to regress vocabulary scores on race (Blacks=0, Whites=1) and year of survey, with and without controlling for the interaction of race and survey year. If the regression coefficient of the interaction term is negative, and the F-test suggests that adding the interaction term contributes significantly to the variance explained, it means that there has been a significant convergence over time between Blacks and Whites in GSS vocabulary scores. In addition, a two-sided Tobit specification should be used to correct for the censored distribution of GSS vocabulary scores which, as displayed in Figure 1, tend to be right-hand censored for Whites. For Blacks, the percentage distribution of vocabulary scores is relatively closer to a normal distribution. When the censored distribution of WORDSUM is not corrected, the size of the Black-White gap in WORDSUM in the 1974-1998 GSS, is 1.43 items. In addition, the size of the Black-White gap in WORDSUM decreases from 1.57 items in 1974 to 1.12 items in 1998. But after correcting for the censored distribution of WORDSUM, the size of the Black-White gap is 1.49 items from 1974 through 1998. Furthermore, the size of the Black-White gap in WORDSUM decreases from 1.66 items in 1974 to 1.16 items in 1998. Thus, the magnitude of

convergence in WORDSUM from 1974 to 1998 between Blacks and Whites is 9.4 percent larger when censoring of the test-score distribution is taken into account. Without correcting for the censored distribution of WORDSUM, Lynn (1998) not only underestimates the size of the Black-White gap in WORDSUM, but also underestimates the magnitude of Black-White convergence over time in WORDSUM.

Table 4 reports the regressions of GSS vocabulary test scores on year of survey and race (Whites=1, Blacks=0), with and without controlling for the interaction of year of survey and race. The regression results presented in Table 4 are categorized by whether or not weighting is applied, whether or not the Tobit specification is used to account for the censored distribution of WORDSUM, and whether or not data from the 1998 GSS are incorporated.[8] The interaction effects are uniformly negative, and they are statistically significant even in two-tailed tests at the 0.05 level in weighted analysis of the full data series, with or without correction for truncation. Contrary to Lynn's conclusion, these findings, too, suggest that the Black-White gap in GSS vocabulary scores narrowed over the survey years from 1974 to 1998.

WHY WEIGHTING MAKES A DIFFERENCE

Why does the narrowing Black-White gap over time in GSS vocabulary scores become statistically significant when data are weighted? The weight variable applied in this analysis contains two components: the number of adults in the household (ADULTS) and the over-sampling of Blacks in 1982 and 1987(OVERSAMP). When data are weighted separately by ADULTS and OVERSAMP,

---

[8] Table A1 in the Appendix is modeled on Table 4, except that year of survey is treated as a series of one-year dummy variables.

as displayed in Table 5, we find that it is the number of adults in the household, not the over-sampling of Blacks in 1982 and 1987, that contributes to the significant Black-White convergence in vocabulary scores. In Table 6, we display the mean and the standard deviation of WORDSUM by survey year with and without applying weights. The narrowing Black-White gap over time in WORDSUM is entirely attributable to the upward trend among Blacks. As demonstrated in Table 6, weighting tends to slightly lower the mean vocabulary scores. This is because respondents who live in larger households (in terms of number of adults living in the household) tend to have lower vocabulary scores.[9] That is, when data are weighted so that samples will be representative of the U.S. population on the individual level, the share of lower scorers increases. As indicated in Table 6, weighting tends to more significantly lower the mean vocabulary scores of Blacks in earlier survey years, such as in 1974, 1976, 1978, and 1982, than it does the scores of Blacks in later years. That is, when data are weighted, the upward WORDSUM trend of Blacks becomes steeper, because the upward trend starts from a lower level. Therefore, when data are weighted, the narrowing Black-White gap in WORDSUM becomes more statistically significant. To test this hypothesis, we examine whether or not the significant effects of the interaction of race and survey year would become insignificant, when data are weighted by ADULTS*OVERSAMP, but no weighting is applied to Blacks in 1974-1978, and only OVERSAMP is used to weight the Black sample in 1982 (because Blacks were oversampled in that year). The regression results reported in the last panel of Table 5 strongly support the hypothesis because the interaction effect is reduced (absolutely) from 0.0176 to 0.0151, and the latter coefficient becomes

---

[9] The correlation between WORDSUM and ADULTS is -.057 ($P < .0001$) for Whites. For Blacks, it is -.065 ($P < .001$).

insignificant.

INTER-COHORT CHANGES IN BLACK-WHITE VOCABULARY TEST SCORE DIFFERENCES

Has the Black-White test score difference been narrowing over time? The best way to answer this question is to examine inter-cohort changes in test score differences between Blacks and Whites (Huang and Hauser, 1998), because the long-term trends in Black-White test score gaps depend on cohort replacement. By analyzing changes in Black-White test score differences across survey years, Lynn (1998) not only fails to make the best use of the data to answer the key research question, but also obscures the significance of the Black-White convergence over time in GSS vocabulary test scores. To be sure, the survey year analysis is not entirely uninteresting, because it reflects current conditions. To capture changes over time in Black-White test score differences, however, cohort analysis is a better strategy. In this section, we extend Huang and Hauser's (1998) intercohort analyses of GSS vocabulary test scores for Blacks and Whites, using data from the GSS, 1974-1998.

In order to conduct the intercohort analysis, several case selections are made. Table 7 shows the details of case selections. The total number of unweighted cases is 12,755 for Whites and 2,308 for Blacks after selections on valid cases. For example, with the White sample, the cumulative sample size in the GSS, 1974-1998, is 31,845, but the ten-item vocabulary test was not administered every year (years administered include: 1974, 1976, 1978, 1982, 1984, 1987 - 1991, 1993, 1994, 1996, and 1998), thus reducing the potential sample size in this analysis to 21,206. Also, due to a switch of rotation design from across surveys to within surveys, one-third of the respondents are “Not Applicable” for all 10 word items after 1988; these “Not Applicable" cases further reduce the potential

sample size in this analysis to 16,379. To construct a variable, CWORDSUM (as we name it to indicate "Corrected WORDSUM") which represents the total number of correct word items, we take those respondents who did not answer all test items, but did answer some test items, to have answered the missing items incorrectly. As it appears in the GSS 1974-1998 codebook, the WORDSUM variable, which purports to measure the total number of correct word items, was constructed in this way; however, there are some coding errors in the construction of the WORDSUM variable.[10] Selection on valid cases of the variable, CWORDSUM, excludes 703 respondents who had "No Answers" on all test items. We include only persons who are at least 20 years old, because many younger people are still in school. Based on Cattell's (Cattell, 1971) hypothesis that "crystallized intelligence" starts to decline at about 60 to 70 years of age and that vocabulary tests, when taken at age 65 or older, may not reflect past verbal ability, we limit our analysis to persons under age 66. Therefore, selection of the age range 20-65 excludes 2,868 cases. Selection on valid cases of the "number of siblings" variable excludes 21 cases; selection on valid cases of the variable "Not living with both own parents at age 16" excludes 4 cases; selection on valid cases of the "number of adults in the household" variable excludes 8 cases; selection on valid respondent's years of education excludes 20 cases. These selections reduce the sample size to12,755 cases for Whites. Going through the same

[10] It is expected that if all word items are "No Answer"s, they should be in the category of WORDSUM = 99. However, 31 cases out of 907 "No Answer" cases in all word items in the entire sample of GSS, 1974-1994, are in WORDSUM = 0. It is also expected that respondents in the category of WORDSUM = 99 should not have answered any item correctly or incorrectly, but should have no answer. However, 24 respondents in WORDSUM = 99 answered WORDH incorrectly, and 22 respondents in WORDSUM = 99 answered WORDH correctly in the entire sample of GSS, 1974-1998. To obtain the correct measure of total number of correct word items, we constructed "CWORDSUM," referring to corrected WORDSUM.

case selections, the sample size is reduced to 2,308 cases for Blacks. The total number of cases is 15,063.

We estimate three models of intercohort trend in verbal ability, based on a two-sided Tobit specification which takes the censored distribution of the GSS vocabulary test scores into account. In the baseline model, we regress sex, race, age, birth cohort, and interaction between race and sex and between race and birth cohort on the total number of correct answers to the 10 CWORDSUM items[11]:

$$E[y] = a + \sum_{1}^{I} ß_i x_i + \sum_{1}^{J} ?_j w_j + d_1 z_1 + d_2 z_2 + d_3 (z_1 z_2) + \sum_{1}^{J} ?_j z_1 w_j, \quad (1)$$

where y is the number of correct CWORDSUM items, a is the intercept, the x_i are dummy variables for age groups, the $ß_i$ are age effects, the w_j are dummy variables for birth cohorts, the $?_j$ are cohort effects, z_1 is a dummy variable for race, z_2 is a dummy variable for sex, the ds are effects of sex and race, and the $?_j$ are effects of race by cohort interactions.

In the GSS, it is possible to distinguish the effects of age from the effects of birth cohort, because the GSS provides repeated measures of the same cohorts at different ages. Nevertheless, it is necessary to assume that there are no period effects on GSS vocabulary test scores due to the fact that year of birth (cohort) is a linear function of survey year (period) and age.[12] Hauser and Huang

---

[11] The number of cases for Whites born in 1909-1919, 1920-1929, 1930-1939, 1940-1949, 1950-1959, 1960-1969, and 1970-1978 are 543, 1,304, 1,876, 2,974, 3,426, 2,107, and 525, respectively. The counts for Blacks were 60, 186, 328, 501, 697, 443, 93, respectively. The small number of cases in the oldest and the youngest cohorts of Blacks suggests that findings about them should be taken with caution.

[12] Year of birth is the difference between survey year and age.

(1997:344-5) investigated the plausibility of this assumption. In the baseline model, we estimate intercohort trends in verbal ability among Blacks and Whites, net of an age effect. The estimated coefficients and their standard errors are reported in Table 8. In the social background model, we add eight social background measures to the baseline model of Equation 1. These measures include father's educational attainment, mother's educational attainment, father's occupational status, number of siblings, nonintact family (at age 16), foreign residence (at age 16), farm background, Southern residence (at age 16), and three dummy variables that flag missing values on father's education, mother's education, and father's occupation. In the education model, we add respondent's years of schooling to the social background model. Years of education are coded into a series of 21 dummy variables with 12 years of education as the reference group.

Figure 2 shows the intercohort trends in GSS vocabulary test scores of Blacks and Whites, as estimated in the baseline model. The Black-White gap in verbal ability has been narrowing.[13] Relative to the earliest birth cohort, the Black-White differences in WORDSUM are 84%, 65%, 62%, 64%, 49%, and 24% as large in successive cohorts.[14] The Black-White convergence is continuous; it shows no sign of ending. Furthermore, about three-fourths of the convergence is attributable to an upward intercohort trend among Blacks, from 4.1 in the earliest cohort to 5.4 in the most recent cohort. About

---

[13] In the baseline model, the likelihood-ratio test statistics for intercohort differences in test scores and for the Race **V** Cohort interactions are both highly significant.

[14] Because of the small number of cases in both the oldest and the youngest cohorts of Blacks, one should take these observations with caution. The Black-White convergence, however, is still highly significant, when the two extreme cohorts (1909-1919 and 1970-1978), either Blacks or Whites, are excluded entirely from the analysis. In the 1960-1969 birth cohort, the Black-White difference is 58.7 percent as large as in the 1920-1929 birth cohort.

one-fourth of the convergence is due to a downward intercohort trend among Whites, from 6.4 in the earliest cohort to 6.0 in the most recent cohort. The significance of the Black-White convergence, the fact that the convergence is continuous, and the fact that most of the convergence is attributable to improvement in Blacks' test scores, are all very encouraging.

To understand why Black-White differences in test scores change over time, it is helpful to examine to what extent the Black-White gap in GSS vocabulary test scores is attributable to Black-White differences in social background and amount of schooling, and to what extent improvements in social background and schooling across time contribute to the Black-White convergence. Figure 3 displays intercohort trends in GSS vocabulary test scores for Blacks and Whites, as estimated in the social background model. After controlling for social background, the upward intercohort trend among Blacks becomes flat. This implies that the observed upward intercohort trend for Blacks is related to an improvement in social background. For Whites, however, improvement in social background across time does not raise test scores correspondingly. A favorable social background for Whites in the later cohorts does not raise their scores to the degree that scores of Whites in the earlier cohorts would be raised if they had the same favorable social background as that of the later cohorts. Therefore, the intercohort trend for Whites is downward after controlling for social background. One alternative explanation is that indicator variables of social background, such as father's years of education and mother's years of education, have increased quantitatively over time, but the selectivity of education has decreased (Wilson and Gove, 1999a, 1999b; Glenn, 1999; Alwin and McCammon, 1999).[15] It could

[15] An education selection effect exists because, given the same education level, recent-borns tend to be less selective with respect to scholastic ability than earlier borns, as the average years of

also be that the quality of education has decreased (Hayes, Wolfer, and Wolfe, 1996). Most importantly, when social background is controlled, Black-White differences are significantly reduced; the Black-White convergence in verbal ability also becomes less rapid, but it remains statistically significant.[16]

Figure 4 shows the intercohort trends in test scores when respondents' years of schooling as well as social background are added to the baseline model. After controlling for both social background and schooling, the Black-White differences in the earliest two birth cohorts are substantially reduced; the differential is more than one correct answer on the WORDSUM test (slightly more than half a standard deviation), and in all but the youngest cohort, the differential is slightly less than one correct answer. The intercohort trends of Blacks and Whites are both downward. The magnitude of the Black-White convergence across birth cohorts, though reduced, continues to be statistically significant, apparently due to the relatively high performance of the youngest cohort of Blacks.[17]

On average, Black-White differences in social background and schooling account for 45% of Black-White differences in GSS vocabulary test scores. That is, changes in social background and schooling contribute to the convergence in test scores. For example, in the baseline model, the change

---

education completed has increased markedly over time.

[16] In the social background model, the likelihood-ratio test statistics for intercohort differences in test scores and for the Race **V** Cohort interactions are both highly significant.

[17] In the education model, the likelihood-ratio test statistics for intercohort differences in test scores are highly significant. For the Race **V** Cohort interactions, the likelihood-ratio test statistics remain to be significant at $p < .005$ level. However, when the most recent cohort, 1970-1978, is excluded from the analysis, the Race **V** Cohort interactions become insignificant.

in Black-White difference between the oldest and the youngest cohorts was 1.8 points; in the social background model, it was 1.5 points; and in the schooling model, it was 1 point. Nevertheless, there are additional signs of convergence in the youngest cohort due to the relatively high performance of the youngest cohort of Blacks. Except for the most recent birth cohort, 1970-1978, cohort changes in social background and schooling account for the narrowing over time of the gap between Blacks and Whites in GSS vocabulary test scores.

CONCLUSION

Using data from the GSS, Lynn (1998) finds a persistent Black-White test score gap over time, but Huang and Hauser (1998) report a significant Black-White test score convergence. While Lynn (1998) analyzes changes in Black-White test score differences across survey year, Huang and Hauser (1998) evaluate changes in Black-White test score differences across birth cohorts. Since the real action about changes over time in Black-White test scores differences is in cohort replacement, cohort analysis is a more sensible strategy. Cohort analysis not only captures the significance of changes over time in Black-White test score differences, but also allows for a Black-White comparison based on a longer time span.

Our analysis of intercohort changes in Black-White vocabulary test score differences suggests that the Black-White test score gap has significantly narrowed over time. The narrowing, in large part, is due to an upward trend among Blacks, whose social background and schooling have improved. The survey year analyses we conducted also show that, when data from the 1998 GSS are incorporated and when data are weighted appropriately, the Black-White test score gap narrowed from 1974 to 1998. Therefore, we cannot agree with Lynn's conclusion that the GSS data provide “no conclusive

evidence that the black-white difference in intelligence has been narrowing over time." No one could claim that the present analysis offers "conclusive" evidence about broader measures of "intelligence" than the vocabulary task measured in the GSS. However, we find that Lynn's analysis of the GSS data is faulty. Proper analyses of the GSS data, either of aggregate trend across survey years or across birth cohorts, provides consistent and significant evidence that vocabulary differences between Black and White Americans narrowed through most of the 20th century.

Bibliography

Alwin, Duane F. and Ryan J. McCammon. 1999. "Aging versus Cohort Interpretations of Intercohort Differences in GSS Vocabulary Scores." *American Sociological Review* 56:625-38.

Davis, James A. and Tom W. Smith. 1992. *The NORC General Social Survey: a User's Guide*. SAGE publications.

Glenn, Norvall 1999. "Further Discussion of the Evidence for an Intercohort Decline in Education-Adjusted Vocabulary." *American Sociological Review* 56:625-38.

Hauser, Robert M., and Min-Hsiung Huang. 1996. "Trends in Black-White Test-Score Differentials." Discussion Paper 1110-96. Madison, Wisconsin: University of Wisconsin-Madison. Institute for Research on Poverty. October.

Hauser, Robert M., and Min-Hsiung Huang. 1997. "Verbal Ability and Socioeconomic Success: A Trend Analysis." *Social Science Research* 26:331-76.

Hayes, Donald, Loreen Wolfer, and Michael Wolfe. 1996. "Schoolbook Simplification and Its Relation to the Decline in SAT-Verbal Scores." *American Educational Research Journal* 33:489-508.

Huang, Min-Hsiung and Robert M. Hauser. 1996. "Trends in Black-White Test-Score Differentials: II. The WORDSUM Vocabulary Test." CDE Working Paper 96-30. Madison, Wisconsin: University of Wisconsin-Madison. Center for Demography and Ecology. November.

Huang, Min-Hsiung and Robert M. Hauser. 1998. "Trends in Black-White Test-Score Differentials: The WORDSUM Vocabulary Test." Pp. 303-32 in *The Rising Curve: Long Term Gains in IQ and Related Measures*, edited by U. Neisser. Washington, DC: American Psychological Association.

Lynn, Richard. 1998. "Has the black-white intelligence difference in the United States been narrowing over time?," *Personality and Individual Differences*, Vol. 25, pages 999-1002.

Maddala, G. S. 1983. *Limited-dependent and Qualitative variables in Econometrics*. Cambridge: Cambridge University Press.

Neisser, Ulric. 1998. *The Rising Curve: Long Term Gains in IQ and Related Measures*. Washington, DC: American Psychological Association.

Wilson, James A. and Walter R. Gove. 1999a. "The Intercohort Decline in Verbal Ability: Does it

Exist?" *American Sociological Review* 64:253-66.

Wilson, James A. and Walter R. Gove. 1999b. "The Age-Period-Cohort Conundrum and Verbal Ability: Empirical Relationships and Their Interpretation" *American Sociological Review* 64:287-302.

Table 1

Differences between Blacks and Whites on Vocabulary Scores Expressed As *d*s, General Social Survey, 1974-1996

		Using Pooled Across-Year SD		Using Pooled Within-Year SD	
Year	Lynn's *d*	*d* (unweighted)	*d* (weighted)	*d* (unweighted)	*d* (weighted)
1974	0.72	0.72	0.73	0.70	0.71
1976	0.79	0.79	0.84	0.77	0.81
1978	0.70	0.70	0.74	0.69	0.71
1982	0.66	0.66	0.69	0.65	0.69
1984	0.67	0.66	0.67	0.65	0.66
1987	0.73	0.72	0.73	0.70	0.71
1988	0.58	0.57	0.58	0.57	0.58
1989	0.65	0.65	0.67	0.64	0.65
1990	0.46	0.46	0.47	0.46	0.48
1991	0.70	0.66	0.67	0.69	0.70
1993	0.56	0.70	0.71	0.74	0.74
1994	0.83	0.75	0.70	0.80	0.74
1996	0.64	0.60	0.61	0.62	0.63
Year and *d*	-0.32 (*p*=.28)	-0.41 (*p*=.17)	-0.55 (*p*=.05)	-0.18 (*p*=.55)	-0.36 (*p*=.23)

Note: The year-specific *d* scores denote the differences between the mean vocabulary scores of whites and the mean vocabulary scores of blacks, divided by either the pooled across-year standard deviation or the pooled within-year standard deviation of vocabulary scores.

Table 2

Differences between Blacks and Whites on Vocabulary Scores Expressed As *d*s, General Social Survey, 1974-1998

		Using Pooled Across-Year SD		Using Pooled Within-Year SD	
Year	Lynn's *d*	*d* (unweighted)	*d* (weighted)	*d* (unweighted)	*d* (weighted)
1974	0.72	0.72	0.73	0.70	0.71
1976	0.79	0.79	0.84	0.77	0.81
1978	0.70	0.71	0.74	0.69	0.71
1982	0.66	0.66	0.69	0.65	0.69
1984	0.67	0.67	0.67	0.65	0.66
1987	0.73	0.72	0.73	0.70	0.71
1988	0.58	0.58	0.58	0.57	0.58
1989	0.65	0.65	0.67	0.64	0.65
1990	0.46	0.46	0.47	0.46	0.48
1991	0.70	0.66	0.68	0.69	0.70
1993	0.56	0.71	0.71	0.74	0.74
1994	0.83	0.75	0.70	0.80	0.74
1996	0.64	0.60	0.61	0.62	0.63
1998	-----	0.56	0.52	0.59	0.55
Year and *d*	-0.32 (*p*=.28)	-0.49 (*p*=.08)	-0.64 (*p*=.01)	-0.26 (*p*=.37)	-0.47 (*p*=.09)

Note: The year-specific *d* scores denote the differences between the mean vocabulary scores of whites and the mean vocabulary scores of blacks, divided by either the pooled across-year standard deviation or the pooled within-year standard deviation of vocabulary scores.

Table 3

Regression of *d* Values on Year of Survey, GSS, 1974-1996, and GSS, 1974-1998

Dependent Variable	*b*	*SE*	R^2
GSS, 1974-1996			
Lynn's *d*	-0.0045	0.0039	0.10
d (unweighted), across-year SD	-0.0049	0.0033	0.16
d (weighted), across-year SD	-0.0069	0.0031	0.30
d (unweighted), within-year SD	-0.0023	0.0037	0.03
d (weighted), within-year SD	-0.0043	0.0033	0.13
GSS, 1974-1998			
d (unweighted), across-year SD	-0.0057	0.0030	0.24
d (weighted), across-year SD	-0.0080	0.0028	0.40
d (unweighted), within-year SD	-0.0030	0.0032	0.07
d (weighted), within-year SD	-0.0055	0.0030	0.22

Note: The year-specific *d* scores denote the differences between the mean vocabulary scores of whites and the mean vocabulary scores of blacks, divided by either the pooled across-year standard deviation or the pooled within-year standard deviation of vocabulary scores.

Table 4

Regression of GSS Vocabulary Test Scores on Year of Survey and Race, With and Without Controlling for the Interaction of Year of Survey and Race, GSS, 1974-1996, and 1974-1998

	Unweighted		Unweighted		Weighted		Weighted	
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
OLS, GSS, 1974-1996								
Intercept	4.045	0.200	3.416	0.565	4.012	0.200	2.994	0.586
Whites (Blacks=0)	1.470	0.045	2.185	0.602	1.463	0.051	2.606	0.620
Year	0.008	0.002	0.015	0.007	0.008	0.002	0.020	0.007
Whites*Year			-0.0083	0.0070			-0.0132	0.0072
Tobit, GSS, 1974-1996								
Intercept	4.090	0.214	3.359	0.602	4.063	0.213	2.936	0.624
Whites (Blacks=0)	1.542	0.049	2.374	0.642	1.530	0.054	2.797	0.661
Year	0.007	0.002	0.016	0.007	0.008	0.002	0.021	0.007
Whites*Year			-0.0096	0.0074			-0.0147	0.0076
OLS, GSS, 1974-1998								
Intercept	3.942	0.184	3.095	0.513	3.968	0.185	2.715	0.536
Whites (Blacks=0)	1.453	0.044	2.419	0.548	1.437	0.049	2.846	0.568
Year	0.010	0.002	0.019	0.006	0.009	0.002	0.023	0.006
Whites*Year			-0.0111	0.0063			-0.0162	0.0065
Tobit, GSS, 1974-1998								
Intercept	3.972	0.197	3.021	0.547	4.004	0.197	2.641	0.571
Whites (Blacks=0)	1.523	0.047	2.609	0.585	1.502	0.052	3.037	0.606
Year	0.009	0.002	0.020	0.006	0.009	0.002	0.024	0.007
Whites*Year			-0.0125	0.0067			-0.0176	0.0069

Note. In order to weight the data, 11 missing cases of ADULTS are deleted. This deletion results in a total of 17,112 cases for the analysis using data from the GSS, 1974-1996. For the analysis using data from the GSS, 1974-1998, this deletion yields a total of 18,337 cases. The weighted and unweighted results are based on the same sample.

Table 5

Regression of GSS Vocabulary Test Scores on Year of Survey, Race, and the Interaction of Year of Survey and Race, GSS, 1974-1998

Model	*b*	*SE*
Tobit, GSS, 1974-1998, unweighted, N=18,337		
Intercept	3.021	0.547
Whites (Blacks=0)	2.609	0.585
Year	0.020	0.006
Whites*Year	-0.0125	0.0067
*Tobit, GSS, 1974-1998, weighted by OVERSAMP*ADULTS, N=18,337*		
Intercept	2.641	0.571
Whites (Blacks=0)	3.037	0.606
Year	0.024	0.007
Whites*Year	-0.0176	0.0069
Tobit, GSS, 1974-1998, weighted by ADULTS, N=18,337		
Intercept	2.641	0.549
Whites (Blacks=0)	3.045	0.585
Year	0.024	0.006
Whites*Year	-0.0175	0.0067
Tobit, GSS, 1974-1998, weighted by OVERSAMP, N=18,337		
Intercept	3.006	0.571
Whites (Blacks=0)	2.615	0.607
Year	0.020	0.006
Whites*Year	-0.0127	0.0069
*Tobit, GSS, 1974-1998, N=18,337, weighted by OVERSAMP*ADULTS for Whites in 1974-1998 and Blacks in years 1984-1998, weighted by OVERSAMP for Blacks in 1982, no weighting applied for Blacks in years 1974-1978*		
Intercept	2.872	0.704
Whites (Blacks=0)	0.022	0.008
Year	2.806	0.732
Whites*Year	-0.0151	0.0082

Table 6
Mean and Standard Deviation of WORDSUM by Year of Survey With and Without Applying Weighting, General Social Survey, 1974-1998

	N		Unweighted Mean		Weighted Mean		Unweighted SD		Weighted SD	
Year	Blacks	Whites	Blacks	Whites	Blacks	Whites	Blacks	Whites	Blacks	Whites
1974	166	1,277	4.66	6.21	4.59	6.17	1.87	2.19	1.80	2.18
1976	116	1,313	4.47	6.19	4.34	6.14	2.09	2.19	2.08	2.16
1978	147	1,324	4.59	6.12	4.56	6.10	1.92	2.21	1.94	2.19
1982	451	1,252	4.71	6.14	4.64	6.12	2.17	2.10	2.16	2.09
1984	155	1,199	4.77	6.20	4.78	6.22	2.00	2.18	1.99	2.16
1987	465	1,166	4.64	6.19	4.60	6.16	2.01	2.16	1.90	2.18
1988	113	779	4.71	5.95	4.68	5.92	1.92	2.18	1.83	2.12
1989	94	840	4.72	6.12	4.62	6.05	2.06	2.15	2.06	2.18
1990	95	732	5.34	6.33	5.30	6.31	2.28	2.11	2.19	2.09
1991	118	817	4.89	6.31	4.83	6.28	1.76	2.06	1.82	2.04
1993	106	857	4.74	6.26	4.73	6.24	1.93	2.02	1.90	2.00
1994	231	1,539	4.78	6.41	4.87	6.37	1.76	1.99	1.78	1.98
1996	253	1,518	4.98	6.27	4.93	6.24	1.95	2.04	1.96	2.03
1998	182	1,043	5.18	6.38	5.19	6.30	1.84	2.02	1.86	2.03

Table 7
Case Selections for Intercohort Analysis of Changes in Vocabulary Test Score Differences between Blacks and Whites: GSS, 1974-1998

Case selection	Number of Cases	
	Whites	Blacks
The cumulative sample size in the GSS, 1974-1998.	31,845	5,183
Vocabulary test was not administered in 1972, 1973, 1975, 1977, 1980, 1983, 1985, and 1986.	21,206	3,759
"Non Applicable" cases due to a switch of rotation design from across surveys to within surveys since 1988.	16,379	3,009
Excluding "No Answer" to all test items cases.	15,676	2,686
Age 20-65.	12,808	2,331
The valid cases of the "number of siblings" variable.	12,787	2,317
The valid cases of the variable "Not living with both own parents at age 16."	12,783	2,314
The valid cases of the "number of adults in the household" variable."	12,775	2,312
The valid cases of the "respondent's years of education completed."	12,755	2,308

Note: Counts are the cumulative numbers of cases remaining after application of the stated criterion.

Table 8

Baseline, Social Background, and Education Models of Vocabulary Test Scores: General Social Survey, 1974-1998

Variable	Baseline		Social background		Education	
	b	*SE*	*b*	*SE*	*b*	*SE*
Intercept	3.69	0.31	3.65	0.29	5.22	0.26
Sex (Men = 1)	0.07	0.11	0.02	0.10	0.06	0.09
Race (White = 1)	2.48	0.31	1.69	0.28	1.35	0.25
Age						
20-29	—	—	—	—	—	—
30-39	0.44	0.05	0.45	0.05	0.24	0.05
40-49	0.56	0.07	0.58	0.06	0.30	0.06
50-59	0.44	0.08	0.46	0.08	0.11	0.07
60-65	0.33	0.11	0.35	0.10	-0.03	0.09
Birth year						
1909-1919	—	—	—	—	—	—
1920-1929	0.22	0.35	-0.16	0.31	-0.47	0.28
1930-1939	0.60	0.33	0.21	0.29	-0.44	0.27
1940-1949	1.03	0.32	0.31	0.29	-0.69	0.26
1950-1959	0.77	0.32	-0.07	0.29	-1.08	0.26
1960-1969	0.89	0.33	-0.34	0.29	-1.28	0.27
1970-1978	1.35	0.38	-0.16	0.34	-0.98	0.31
Interaction of sex and race	-0.29	0.11	-0.24	0.10	-0.39	0.09
Interaction of birth year and race						
1909-1919	—	—	—	—	—	—
1920-1929	-0.38	0.36	-0.16	0.32	0.01	0.29
1930-1939	-0.82	0.34	-0.77	0.30	-0.41	0.28
1940-1949	-0.89	0.33	-0.84	0.29	-0.28	0.27
1950-1959	-0.85	0.32	-0.89	0.29	-0.32	0.26
1960-1969	-1.19	0.33	-0.98	0.30	-0.44	0.27
1970-1978	-1.79	0.39	-1.47	0.35	-1.00	0.32
Father's education			0.06	0.01	0.02	0.01
Mother's education			0.10	0.01	0.04	0.01
Father's occupational status			0.01	0.00	0.01	0.00
Number of siblings			-0.08	0.01	-0.04	0.01

Table 8 (Continued)

Not living with both parents at age 16	0.02	0.05	0.10	0.05
Lived in foreign country at age 16	0.13	0.33	-0.17	0.30
Farm background	-0.45	0.04	-0.32	0.04
Lived in the South at age 16	-0.46	0.04	-0.35	0.03
Missing father's education	-0.45	0.06	-0.09	0.05
Missing mother's education	-0.67	0.06	-0.27	0.05
Missing father's occupational status	0.19	0.07	0.03	0.06
No schooling			-0.12	0.68
Years of school				
1			-3.17	0.98
2			-1.42	0.66
3			-2.58	0.35
4			-2.61	0.32
5			-2.37	0.26
6			-1.51	0.20
7			-1.57	0.14
8			-1.28	0.09
9			-1.09	0.09
10			-0.73	0.07
11			-0.71	0.07
12			—	—
13			0.38	0.06
14			0.72	0.05
15			1.03	0.07
16			1.64	0.05
17			1.84	0.09
18			2.03	0.09
19			2.44	0.14
20+			2.42	0.12

Note. Entries for reference groups in sets of more than two dummy variables are marked by a dash. Thus, main effects of race pertain to the Black-White difference in the 1909-1919 cohort; main effects of cohorts pertain to contrasts among Blacks, relative to the 1909-1919 cohort; and Race **V** Cohort interactions pertain to cohort-specific Black-White differences. Age effects are expressed relative to ages 20-29, and education effects are expressed relative to the completion of 12 years of school.

APPENDIX

Table A1

Regression of GSS Vocabulary Test Scores on Year of Survey Dummies and Race With and Without Controlling for the Interaction of Year of Survey (Continuous) and Race, GSS, 1974-1996, and GSS, 1974-1998

	Unweighted		Unweighted		Weighted		Weighted	
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
OLS, GSS, 1974-1996								
Intercept	4.737	0.069	4.653	0.103	4.698	0.069	4.556	0.105
Whites (Blacks=0)	1.460	0.046	2.127	0.605	1.465	0.051	2.588	0.620
Year (Dummies)								
1974								
1976	-0.031	0.079	-0.020	0.079	-0.058	0.075	-0.039	0.076
1978	-0.083	0.078	-0.057	0.082	-0.063	0.076	-0.018	0.080
1982	-0.049	0.076	0.010	0.093	-0.056	0.073	0.037	0.089
1984	0.009	0.080	0.078	0.101	0.056	0.078	0.172	0.100
1987	-0.032	0.076	0.056	0.110	-0.012	0.073	0.138	0.111
1988	-0.219	0.090	-0.124	0.125	-0.216	0.089	-0.054	0.126
1989	-0.069	0.088	0.034	0.128	-0.105	0.086	0.070	0.129
1990	0.188	0.092	0.298	0.135	0.195	0.090	0.381	0.136
1991	0.121	0.088	0.237	0.137	0.116	0.088	0.312	0.139
1993	0.057	0.088	0.187	0.147	0.075	0.086	0.296	0.149
1994	0.188	0.075	0.324	0.144	0.202	0.073	0.432	0.146
1996	0.095	0.075	0.244	0.154	0.097	0.073	0.349	0.157
Continuous Year*Whites			-0.0077	0.0070			-0.0130	0.0072

Table A1. Continued

	Unweighted		Unweighted		Weighted		Weighted	
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
Tobit, GSS, 1974-1996								
Intercept	4.742	0.074	4.645	0.110	4.705	0.074	4.547	0.112
Whites (Blacks=0)	1.530	0.049	2.305	0.646	1.531	0.054	2.776	0.661
Year (Dummies)								
1974								
1976	-0.035	0.084	-0.022	0.085	-0.068	0.080	-0.047	0.081
1978	-0.090	0.084	-0.059	0.087	-0.070	0.081	-0.020	0.085
1982	-0.068	0.081	0.002	0.099	-0.076	0.078	0.026	0.095
1984	-0.006	0.085	0.074	0.108	0.042	0.083	0.171	0.107
1987	-0.053	0.082	0.049	0.118	-0.037	0.078	0.130	0.118
1988	-0.251	0.096	-0.140	0.133	-0.249	0.095	-0.070	0.134
1989	-0.099	0.094	0.020	0.137	-0.140	0.092	0.053	0.137
1990	0.190	0.098	0.317	0.144	0.192	0.096	0.397	0.145
1991	0.109	0.094	0.243	0.146	0.100	0.094	0.317	0.148
1993	0.034	0.094	0.185	0.157	0.050	0.092	0.294	0.159
1994	0.178	0.080	0.336	0.153	0.188	0.079	0.443	0.156
1996	0.076	0.080	0.248	0.164	0.077	0.079	0.356	0.167
Continuous Year*Whites			-0.0090	0.0075			-0.0144	0.0076

Table A1. Continued

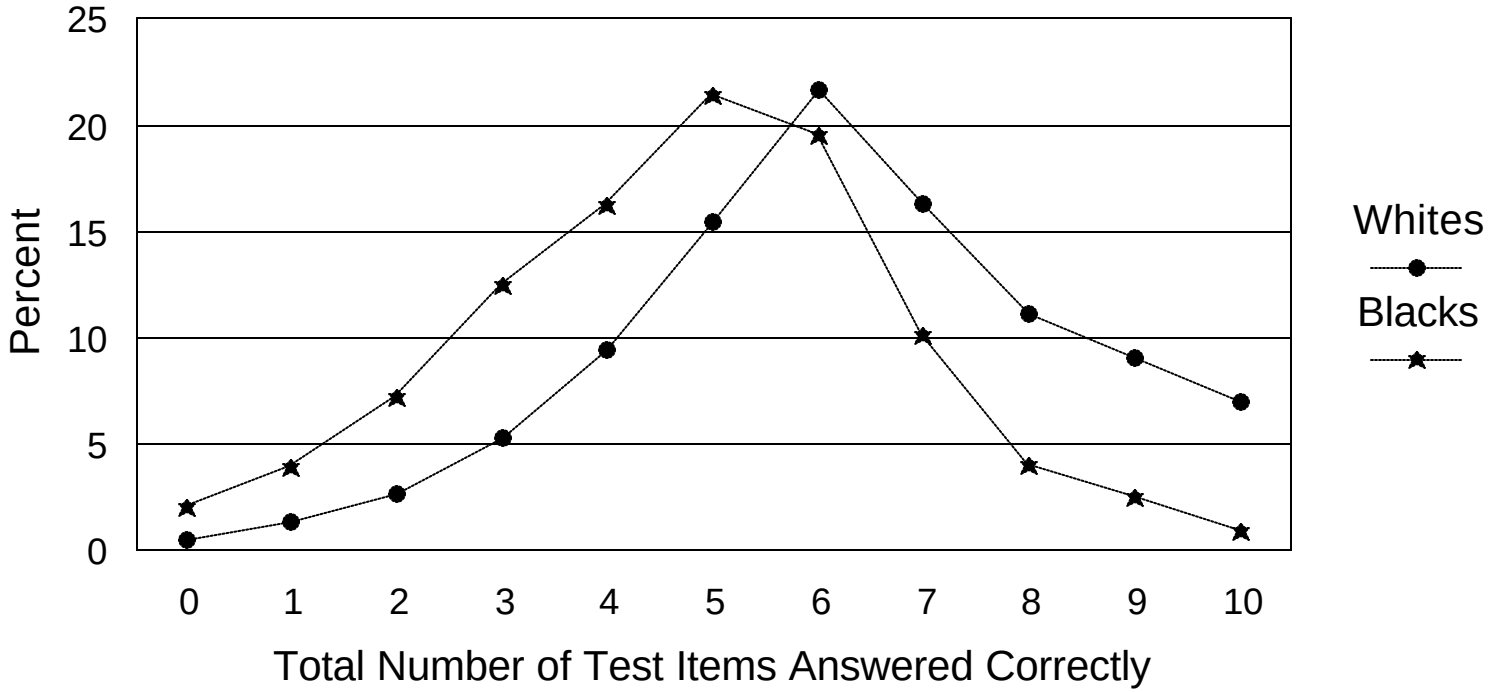
	Unweighted		Unweighted		Weighted		Weighted	
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
OLS, GSS, 1974-1998								
Intercept	4.753	0.068	4.634	0.100	4.722	0.068	4.536	0.102
Whites (Blacks=0)	1.442	0.044	2.332	0.552	1.438	0.049	2.820	0.569
Year (Dummies)								
1974								
1976	-0.031	0.078	-0.016	0.079	-0.057	0.075	-0.034	0.075
1978	-0.083	0.078	-0.048	0.081	-0.063	0.075	-0.008	0.078
1982	-0.052	0.075	0.028	0.090	-0.055	0.072	0.057	0.086
1984	0.009	0.079	0.100	0.097	0.056	0.077	0.197	0.096
1987	-0.035	0.076	0.083	0.105	-0.013	0.073	0.171	0.105
1988	-0.220	0.089	-0.093	0.119	-0.216	0.088	-0.019	0.120
1989	-0.069	0.088	0.067	0.122	-0.104	0.086	0.108	0.122
1990	0.188	0.091	0.333	0.128	0.195	0.090	0.421	0.129
1991	0.121	0.088	0.274	0.129	0.116	0.087	0.355	0.131
1993	0.057	0.087	0.229	0.138	0.075	0.086	0.344	0.140
1994	0.188	0.074	0.368	0.134	0.201	0.073	0.482	0.136
1996	0.095	0.074	0.291	0.142	0.096	0.073	0.403	0.146
1998	0.223	0.081	0.436	0.155	0.189	0.081	0.522	0.159
Continuous Year*Whites			-0.0102	0.0063			-0.0159	0.0065

Table A1. Continued

	Unweighted		Unweighted		Weighted		Weighted	
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
Tobit, GSS, 1974-1998								
Intercept	4.759	0.072	4.626	0.107	4.729	0.073	4.527	0.109
Whites (Blacks=0)	1.511	0.047	2.508	0.589	1.504	0.052	3.006	0.606
Year (Dummies)								
1974								
1976	-0.034	0.084	-0.018	0.084	-0.067	0.080	-0.042	0.081
1978	-0.090	0.083	-0.051	0.086	-0.070	0.080	-0.010	0.084
1982	-0.070	0.081	0.019	0.096	-0.075	0.077	0.046	0.092
1984	-0.006	0.085	0.095	0.104	0.042	0.083	0.196	0.103
1987	-0.056	0.081	0.075	0.112	-0.037	0.078	0.162	0.112
1988	-0.251	0.095	-0.110	0.127	-0.250	0.094	-0.035	0.128
1989	-0.099	0.094	0.053	0.130	-0.140	0.091	0.091	0.130
1990	0.190	0.098	0.352	0.137	0.192	0.096	0.437	0.138
1991	0.109	0.094	0.280	0.138	0.099	0.093	0.359	0.140
1993	0.035	0.093	0.227	0.147	0.050	0.092	0.342	0.149
1994	0.178	0.080	0.379	0.143	0.188	0.078	0.492	0.145
1996	0.076	0.080	0.296	0.152	0.076	0.078	0.410	0.155
1998	0.215	0.087	0.454	0.165	0.179	0.087	0.541	0.169
Continuous Year*Whites			-0.0114	0.0067			-0.0172	0.0069

Note. In order to weight the data, 11 missing cases of ADULTS are deleted. This deletion results in a total of 17,112 cases for the analysis using data from the GSS, 1974-1996. For the analysis using data from the GSS, 1974-1998, this deletion yields a total of 18,337 cases. The weighted and unweighted results are based on the same sample.

Figure 1
Percentage Distribution of the 10-Item Vocabulary Test Scores
Ages 24-89, General Social Survey, 1974-1998



There are 15,656 whites and 2,692 Blacks in the sample.

Figure 2
The Inter-cohort Trends in Vocabulary Test Scores for Blacks and Whites
Predicted by TOBIT Baseline Model: GSS, 1974-1998

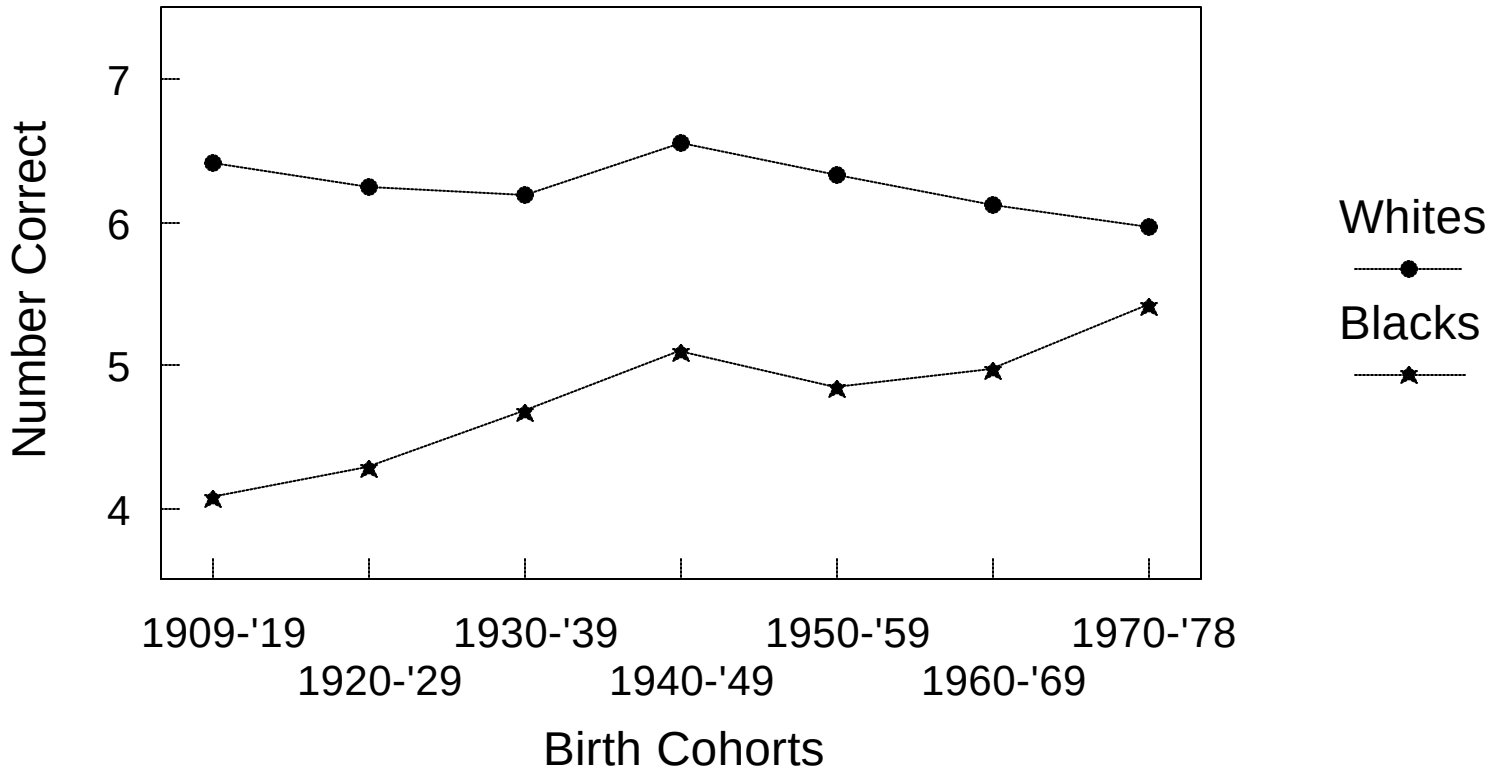


Figure 3
The Inter-cohort Trends in Vocabulary Test Scores for Blacks and Whites
Predicted by TOBIT Social Background Model: GSS, 1974-1998

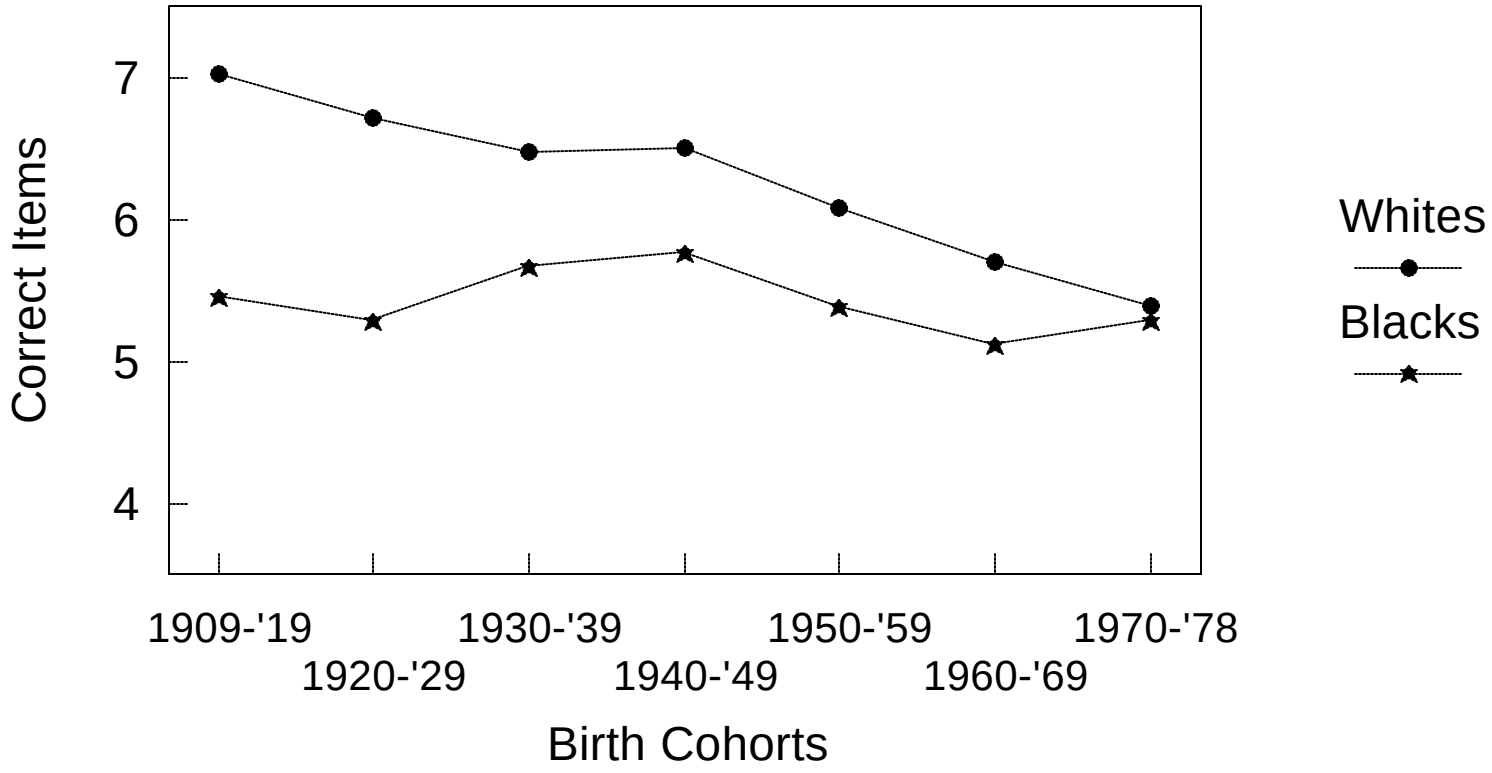


Figure 4
The Inter-cohort Trends in Vocabulary Test Scores for Blacks and Whites
Predicted by TOBIT Education Model: GSS, 1974-1998