

Bacharach, V. R., Baumeister, A. A., & Furr, R. M. (2003). Racial and gender science achievement gaps in secondary education. *Journal of Genetic Psychology*, 164(1): 115-126. (March 2003) Published by Taylor & Francis (ISSN: 1940-0896).

Racial and gender science achievement gaps in secondary education

Verne R. Bacharach, Alfred A. Baumeister, and R. Michael Furr

ABSTRACT

A substantial disparity exists for academic achievement in science between Black and White primary-school children. A similar gap exists between boys and girls. The extent to which secondary education influences these achievement gaps has not been established. The authors report analyses showing how these science achievement gaps change as a function of secondary education. Analyses of data from a large, nationally representative longitudinal study of academic achievement showed that racial disparities and disparities associated with gender continue to increase throughout high school.

A large and diverse literature attests to the existence of a significant and enduring academic achievement gap between Black and White children (Bankston & Caldas, 1997; Entwistle & Alexander, 1988; Grissmer, Flanagan, & Williamson, 1998; Loehlin, Lindzey, & Spuhler, 1975; Myerson, Rank, Raines, & Schnitzler, 1998). Not only does the disparity occur early--Black children enter primary school at an academic disadvantage (National Center for Education Statistics [NCES], 2000)--but also the difference in academic achievement between Black and White children increases during elementary school (Caldas, 1993; Phillips, Crouse, & Ralph, 1998). Uncertainty remains as to whether this trend continues during secondary school. Moreover, it is not known if the size of the changes in achievement gaps during high school is similar across academic disciplines. Although changes in the achievement gap in mathematics and reading during secondary school have been investigated, questions persist as to the magnitude and change in size of the science achievement gap during secondary school.

Cross-sectional data from NCES (1999a) indicate only modest success in efforts designed to reduce the academic achievement differential. That is, secondary education may exert a small positive effect in reducing the educational achievement gap in selected academic areas. These benefits, however, are questionable because cross-sectional data, such as those reported by the NCES, do not address issues concerning the role and the magnitude of effects specifically associated with secondary education. In particular: Is secondary education a source of the reduction in the Black--White educational achievement gap or do cohort (generational) effects confound the comparisons?

Differences in event dropout rates for Black and White secondary school children are one source of a relevant confounding cohort effect. Since the 1970s, event dropout rates have been declining faster for Blacks than for Whites. In the case of the NCES (1999b) data, this particular type of cohort effect might act to increase the Black--White achievement

difference on the assumption that the more precipitous decline in Black dropout rates is likely to increase the number of academically weak students in a Black sample. Dropout rates are one example of a large number of variables that might be operating to produce cohort effects.

LONGITUDINAL STUDIES

To determine whether secondary education is having the desired effect of reducing the educational achievement gap between Black and White secondary-school students, it is necessary to examine achievement change data from longitudinal studies. A review of relevant published longitudinal studies revealed three problems that compromise conclusions concerning the effects of secondary education on the Black--White achievement disparity.

First, reported data come from studies involving nonrepresentative samples of children and schools. For example, Project 2000 was a well-executed, 4-year longitudinal study of high school students who scored between the 25th and 65th percentiles on a standardized achievement test that was administered when these children were freshmen. However, the external validity of these results was compromised because all of the children in this study

attended school in one California school district (Nyberg, McMillin, O'Neill-Rood, & Florence, 1997).

Second, other reports were based on data obtained before or shortly after federal government initiatives to fund and evaluate education programs and often do not provide information regarding changes in academic achievement during high school. Examples of these projects include Project TALENT (Shaycoft, 1967), the Explorations in Equality of Opportunity Study (Alexander, Eckland, & Griffin, 1975), the Wisconsin Study (Sewell, Hailer, & Portes, 1969), and the Educational Testing Service's Growth Study (Hilton, Beaton, & Bower, 1971). Because of the numerous educational reforms that were put in place in the late 1970s and early 1980s, these studies have limited value for evaluations of the effects of current educational practices on achievement gaps.

Third, a number of reported longitudinal studies were narrowly designed to evaluate effects of specific educational intervention programs on school achievement (Reynolds & Temple, 1998). Although these studies provide useful information concerning the effectiveness of particular programs specially designed to reduce the racial achievement gap, they do not reveal the extent to which racial achievement inequities are influenced by secondary education in a majority of U.S. schools that do not offer these intervention programs.

We found only three sources of nationally representative longitudinal data regarding academic achievement in secondary schools. Two sources are from longitudinal studies sponsored by the U.S. government. Each involves large representative samples of high school students.

One of these studies, the High School and Beyond study (HS&B; Zahs, Pedlow, Morrissey, Marnell, & Nichols, 1995), and its auxiliary companion, the High School Transcript study, followed academic progress of a cohort of sophomores through high school. A second study, the National Longitudinal Survey of Labor Market Experience of Youth and its subset, Children of the National Longitudinal Survey of Youth (NLSY), is a widely reported source of longitudinal data regarding academic achievement (Pergamit, 1995). The NLSY and the HS&B studies were begun in 1979 and 1980, respectively, making each data set about 20 years old. In addition, neither survey includes longitudinal data on science achievement.

The most recent longitudinal study of academic achievement in secondary school (NCES, 1996) is the National Educational Longitudinal Study (1988; NELS: 88/94). The study began with a representative sample of 8th graders from a representative sample of U.S. middle schools. The participants' academic progress was followed through high school. We used these data in the present analyses to examine change in the racial academic achievement gap in science from 8th grade through 12th grade.

No longitudinal study that allows direct comparisons of the academic achievement gap across the entire range of school years has been reported. However, Phillips and colleagues (1998) attempted to solve this problem by merging the results of two separate longitudinal studies: (a) Prospects: The Congressionally Mandated Study of Educational Growth and Opportunity, which followed a cohort from 1st through 9th grade, and (b) data from NELS: 8 8/94. The researchers showed that Black children tend to fall further behind White children in reading achievement but

not in mathematics as they progress through high school. They also did not report on science achievement changes that might occur during secondary school.

Disparities in science achievement also exist between boys and girls. In general boys tend to do better on science achievement assessments (NCES, 1999c). As is the case with the racial achievement gap in science, the gender achievement gap appears during primary school, and cross-sectional data seem to indicate that the gap persists through secondary school. Our concern is with the change in the science achievement gender gap among Black and White children as they progress through secondary school.

No effort was made to isolate causal factors producing the initial racial and gender disparities or causing changes in those disparities. Rather, the intent was to construct an accurate and informative statistical picture of science achievement disparities and of changes in those disparities with secondary schooling.

Because our investigation was concerned with the changes in the racial and gender science achievement gaps as a function of secondary school experience, only those children who successfully completed primary and secondary education were included. The focus of this study was on the restricted sample for two reasons: (a) the restricted sample is representative of the large majority of U.S. school children, approximately 88% of whom eventually graduate from high school (NCES, 1998); and (b) it would make little sense to inquire about the effects of schooling using data from children who had not successfully completed their education. Interpretation of the results of these analyses must be conditioned by the characteristics of the sample population.

METHOD

Participants

Participants were children enrolled in The National Education Longitudinal Study: 88/94. Data from the NELS: 88/94 (NCES, 1996) were used because this research is the most recent longitudinal study of educational achievement involving a nationally representative sample of secondary school children. We used a sampling procedure designed to obtain a nationally representative sample of 8th-grade students from a nationally representative sample of middle schools.

TABLE 1
Zero-Order Correlations for Science Achievement
Item Response Theory Test Scores

Variable	8th grade	10th grade	12th grade
8th grade		.72	.68
10th grade			.79
Race	.21	.23	.21
Gender	.10	.16	.17

Note. For the dummy code, Black = 0 and female = 0. All *r* values are statistically significant (*p* < .05).

Approximately 26,000 8th-grade children from 1,052 schools were initially selected in 1988 for the present study. Funding restrictions reduced the effective sample size to approximately 21,000 children. In 1994, a postsecondary education follow-up was conducted on a sample of 14,915 participants, 13,120 of whom were in the original sample and had participated in each of two previous follow-up surveys conducted 2 years (10th grade) and 4 years (12th grade) after the initial survey. Of these children, 11,617 graduated from high school in 1992 with their 8th-grade cohort. The percentage of White children (84%) in this graduation cohort was larger than the percentage of Black children (72%). The present analyses of the NELS: 88/94 data were based on a subset of these children.

The sample consisted of 668 Black children and 5,463 White children for whom there was a complete data set that included the following variables: (a) gender; (b) race; and (c) scores on science achievement tests given during the 8th-, 10th-, and 12th-grade years. Fifty-two percent of the Black children and 51% of the White children were girls.

Achievement Test

Children who participated in the NELS: 88/94 study were administered academic achievement tests covering several content areas. Our interest focused on achievement in science. The Educational Testing Service constructed the science achievement test for the NELS: 88/94 survey. The 25-item test covered topics in life, earth, and physical sciences. Children were given a fixed time limit for completing the test, but the test was not considered speeded.

The NELS: 88/94 data set includes item response theory science achievement test scores (IRT scores). We calculated IRT scores by examining the pattern of correct and incorrect responses to items on the science achievement test. We then used item difficulty and discriminability to construct a score scale common to test results across testing intervals (Crocker & Algina, 1986; NCES, 1994; Schultz & Nicewander, 1997).

These values have at least two advantages over raw score values. In theory, IRT scores are less likely than raw scores to be influenced by variables such as guessing and, more important for our purpose, they allow vertical scaling of test scores across grade level. The NCES included these values in the NELS: 88/94 data set for the express purpose of longitudinal analyses.

Data Analysis

As previously noted, the primary purpose of the present study was to examine and characterize changes in the size of the academic achievement gap between Black and White, and male and female students as they progress through secondary school. For the analyses, we used longitudinal data from a cohort of students tested at three different time periods. We used hierarchical linear modeling (HLM) to treat these data. HLM involves two levels of analysis. The procedure first estimates slopes (regression coefficients) and intercepts for each student. The slopes represent growth curves or average yearly change in test scores across time. Intercepts are estimates of 8th-grade achievement test scores. These values are then used as dependent variables to estimate average slopes and intercepts for groups of students.

If science achievement scores improve at the same rate for all groups, each group will have a similar slope. This outcome would suggest that the achievement gap does not change during secondary school. If there are significant slope differences between groups, the pattern of these differences can be examined to determine if achievement gaps are increasing or decreasing over time.

RESULTS

Table 1 presents zero-order correlations among the variables. Means and standard deviations for the quantitative variables of interest are provided in Table 2.

TABLE 2
Factor Loadings for the Learning Process Questionnaire Subscales

Race	Males		Females		Total	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>8th grade</i>						
Black	17.1	4.0	16.07	3.8	16.9	3.9
White	20.7	5.0	19.7	4.3	20.2	4.7
Mean	20.3	5.0	19.4	4.4		
<i>10th grade</i>						
Black	18.7	4.9	18.1	4.5	18.4	4.7
White	24.3	5.8	22.5	5.3	23.4	5.6
Mean	23.7	6.0	22.0	5.4		
<i>12th grade</i>						
Black	20.2	5.5	19.1	5.2	19.6	5.3
White	26.0	5.9	24.0	5.6	25.0	5.8
Mean	25.4	6.1	23.4	5.7		

Results of HLM Analyses

Table 3 shows the results of the HLM analysis for main effects of race and gender on science achievement IRT scores. Black students were the reference group in the analysis-of-race differences, and girls were the reference group in the analysis-of-gender differences.

TABLE 3
Results of Hierarchical Linear Modeling Analyses: Main Effects Models

Outcome	Race		Gender	
	Intercept ^a	B	Intercept ^b	B
8th-grade achievement	16.92***	3.55***	19.57***	1.05***
SD units		0.75		0.22
Change per year	0.68***	0.51***	1.01***	0.26**
SD units	0.14	0.11	0.21	0.05

Note. B = regression weight. All *dfs* = 6, 129. SD = slope difference.

^aThe intercept values are the estimated mean achievement score and yearly change for Black students.

^bThe intercept values are the estimated mean achievement score and yearly change for female students.

p* < .01. *p* < .001.

Black and White students had significantly different average 8th-grade science achievement test scores. More important, the average yearly increase in achievement for White students was greater than the average yearly increase for Black students. Specifically, in the 8th grade, Black students had an estimated average IRT science achievement score of 16.92, whereas White students had an estimated average score that was 3.55 IRT points higher (*p* < .001). Following the 8th grade, Black students showed a significant average yearly increase in science achievement scores of 0.68 IRT points (*p* < .001) compared with an average yearly increase for White students that was approximately 0.51 points higher than the average yearly increase for Black students (*p* < .001). Thus, Black students finished the 8th grade with lower science achievement scores than White students, and the size of this disparity continued to increase during secondary school.

Average 8th-grade science achievement scores were significantly different for female and male students. More important, the average yearly increase for boys was larger than the average yearly increase for girls. In 8th grade, girls had an estimated average science IRT achievement score of 19.57; boys had an estimated average score that was 1.05 points higher (*p* < .001). Following 8th grade, girls showed a significant average yearly increase in achievement scores of 1.01 IRT points per year (*p* < .001), whereas boys showed a yearly increase that was approximately 0.26 points higher than the increase for girls (*p* < .001). Thus, girls finished the 8th grade with lower science achievement scores than boys, and the size of this disparity continued to increase during secondary school.

When the full factorial model was examined, there was a race-by-gender interaction for mean 8th-grade IRT scores (*p* < .001); however, there was no race-by-gender interaction associated with average yearly rates of change (*p* > .05). Figure 1 illustrates the change in IRT science achievement scores across grade levels for each group in the study.

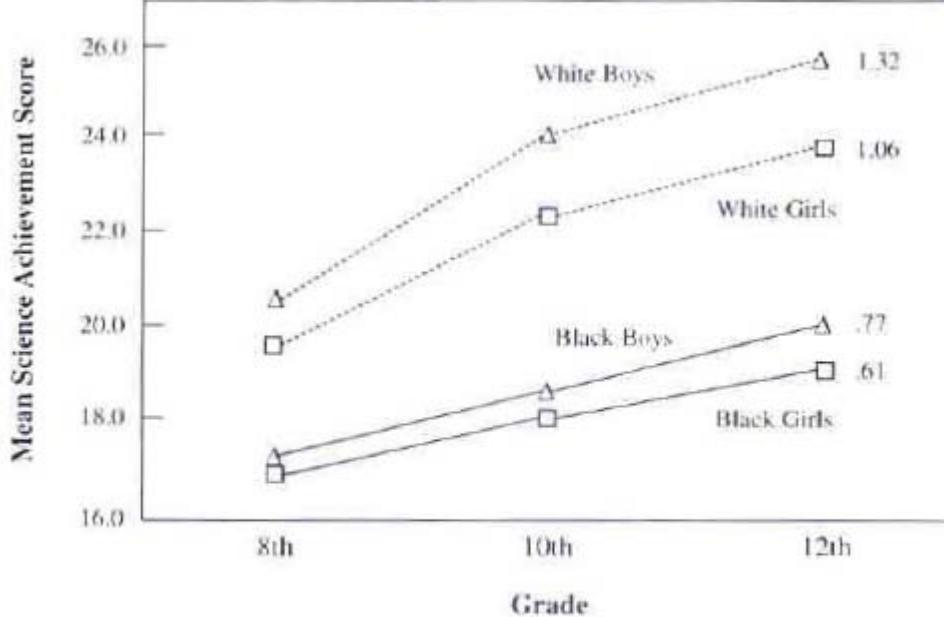
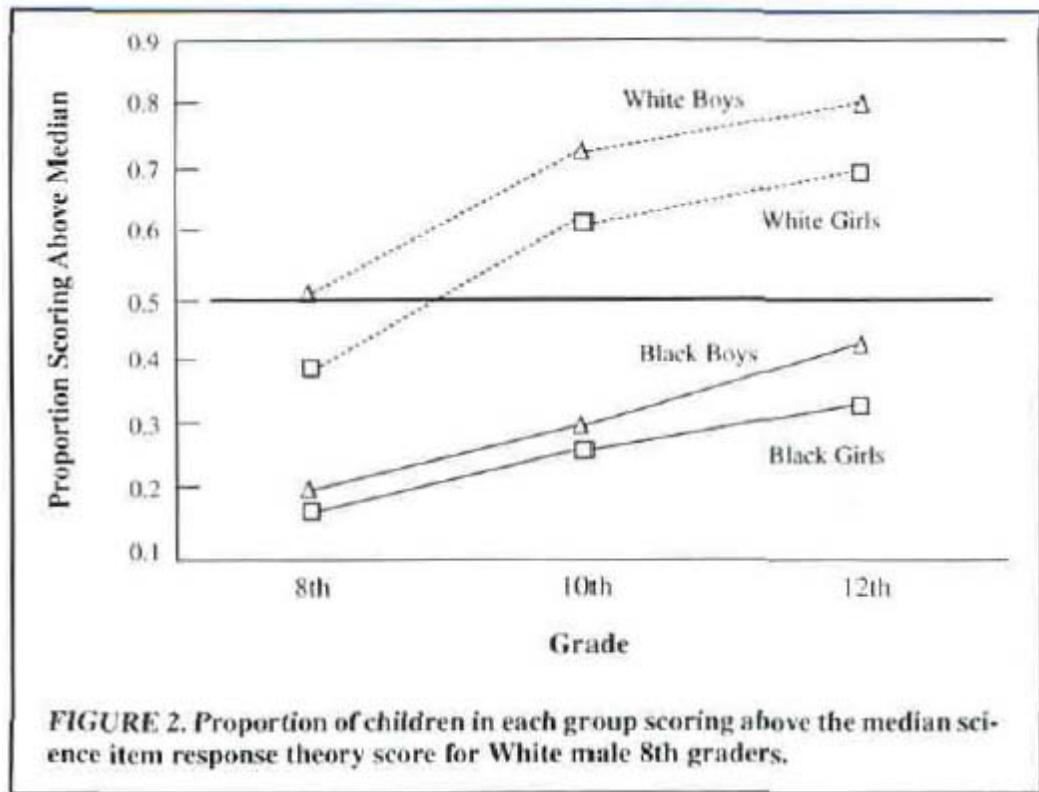


FIGURE 1. Mean item response theory science achievement scores for groups across grades. Estimated slopes for each group are included. All slope differences are statistically significant, $p < .05$.

The size of the racial achievement gap and the increase in the size of this disparity may seem small, but in fact these findings have important educational consequences. One way to illustrate the practical importance of this gap and the increase in the size of the gap is to examine the proportion of children in each of the target groups who scored above the median IRT science score of 8th-grade White males. Figure 2 presents the proportion of children scoring above the White male 8th-grade baseline at each grade level.

Figure 2 shows a way of estimating the magnitude of the gap size in terms that can be readily appreciated. For example, it can be seen that fewer than 50% of Black 12th graders were performing at a level comparable to the average test performance of 8th-grade White boys. Taken together, Figures 1 and 2 also illustrate the relative magnitude of the gender and race effects. Although test performance of boys was superior to that of girls, this difference was small compared with racial differences.



DISCUSSION

In the present study, we focused on a representative group of Black and White children who had successfully completed their secondary school education. We found a substantial academic achievement gap between Black and White children and between boys and girls prior to secondary school. Secondary schooling did not reduce or compensate for the achievement differences that developed during primary school. In fact, the science academic achievement gaps between Black and White children and between boys and girls continued to grow during high school.

These analyses represent a conservative estimate of the size and growth of the academic achievement gap between races during secondary school. Children in the present study could be considered successful students in that they completed high school at the expected time. In addition, the average grade point average for the 8th-grade children in the present study was approximately 3.0 on a 4-point scale. If the study had used a cross-sectional sample of all children in a particular grade at any one point in time, perhaps the racial achievement gap would have been even larger than the one reported in the present study. The gap may have been larger simply because a cross-sectional sample probably would have included academically less talented students such as those who do not complete their high school education or who are late in completing it.

Why does the racial academic achievement disparity in science exist and continue to grow throughout primary and secondary school? Although this question has been frequently addressed, no definitive conclusion has emerged. Some researchers have argued that the schools attended by Black children are of poorer quality than the schools attended by White children. Others have suggested that Black children are less likely than White children to value education (Osbourne, 1995, 1997; Steele, 1997). According to Steele (1997), racial stigma reduces the self-esteem of Black children, leading to disidentification with education and poor achievement. Various family factors favoring the average White child have been offered as explanations for relatively poor academic achievement among Black children (see Booth & Dunn, 1996; Finn & Owings, 1994). De facto segregation has been nominated for blame (Bankston & Caldas, 1996; Rumberger & Willms, 1992). Academic tracking has received its share of attention as a factor contributing to the differential (Davenport et al., 1998; Nyberg et al., 1997). The list of factors that might account for the achievement gap is inclusive. No single study or analysis is likely to provide a definitive answer.

Phillips et al. (1998), in their analyses of mathematics and reading IRT test scores, found that the reading achievement gap between Black and White children continued to grow during secondary school but that the mathematics achievement disparity between these children did not increase. We found that the growth of science achievement disparities varied depending on gender and race. Taken together, the results of these analyses suggest that none of the main effects theories previously discussed are likely to capture the details associated with the range of secondary school achievement gap outcomes found across studies. Theories that appeal to social and economic disparities between the races fail to account for the gender differences within races or for different outcomes depending on the type of achievement under evaluation. A similar complaint could be made for main effect theories that attempt to account for achievement disparities associated with gender.

Regrettably, secondary education does not contribute to a reduction in the science achievement gaps associated with race and gender. It would be incorrect, however, to assume that this failure is caused by poor-quality secondary education. There is no adequate baseline control against which the achievement gaps can be compared. It is possible, for example, that educational practices that differ from current ones could lead to achievement-gap increases larger than the ones associated with current practices. It also is possible that the gaps and the growth of the gaps are under control of factors, such as societal factors, that are not influenced by educational practices.

Analytic Qualifications

As is the case with many longitudinal studies, the NELS: 88/94 data are compromised by participant attrition. In this study, attrition was minimal (Phillips et al., 1998). About 90% of the base-year panel participated in the first and second follow-up studies. The panel completion rates were in excess of 90% for both Black and White children (NCES, 1994). Given these relatively low attrition rates and the large sample size, effect size estimates reported here are probably stable.

One also must consider the nature of the achievement test scores used in the present and in previous analyses (Phillips et al., 1998; Ralph & Grouse, 1997) when interpreting NELS: 88/94 longitudinal data. Some have suggested that IRT scores may be subject to various types of scaling problems (Schulz & Nicewander, 1997). The extent to which these problems actually exist and the extent to which they might influence estimations reported here are not clear. Furthermore, national data based on alternative scaling and testing methods are not available. Future research that relies on more adequate testing and scaling procedures might conceivably produce results that differ from those reported in this study; however, we are confident that the NELS: 88/94 data represent the best information available concerning the secondary school science achievement gap between Black and White, and male and female children.

REFERENCES

- Alexander, K. L., Eckland, B. K., & Griffin, L. J. (1975). The Wisconsin model of socioeconomic achievement: A replication. *American Journal of Sociology*, 81, 324-342.
- Bankston, C. L., & Caldas, S. J. (1996). Majority African American schools and social injustice: The influence of defacto segregation on academic achievement. *Social Forces*, 75, 535-555.
- Bankston, C. L., & Caldas, S. J. (1997). The American school dilemma: Race and scholastic performance. *The Sociological Quarterly*, 38, 423-429.
- Booth, A., & Dunn, J. F. (1996). Family-school links. Mahwah, NJ: Erlbaum.
- Caldas, S. J. (1993). Reexamination of input and process factor effects on public school achievement. *Journal of Educational Research*, 86, 206-214.
- Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart and Winston.
- Davenport, E. C., Davison, M. L., Kuang, H., Ding, S., Kim, S. K., & Kwak, N. (1998). High school mathematics course-taking by gender and ethnicity. *American Educational Research Journal*, 35, 497-514.
- Entwistle, D. R., & Alexander, K. L. (1988). Factors affecting achievement test scores and marks of Black and White first graders. *Elementary School Journal*, 88, 449-471.
- Finn, J. D., & Owings, M. F. (1994). Family structure and school performance in the eighth grade. *Journal of Research and Development in Education*, 27, 176-187.
- Grissmer, D., Flanagan, A., & Williamson, S. (1998). Why did the Black-White score gap narrow in the 1970s and 1980s? In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 182-226). Washington, DC: Brookings Institute Press.
- Hilton, T. L., Beaton, A. E., & Bower, C. P. (1971). Stability and instability in academic growth: A compilation of longitudinal data. Princeton, NJ: Educational Testing Service.

- Loehlin, J. C., Lindzey, G., & Spuhler, J. N. (1975). Race differences in intelligence. San Francisco: Freeman.
- Myerson, J., Rank, M. R., Raines, F. Q., & Schnitzler, M. A. (1998). Race and general cognitive ability: The myth of diminishing returns to education. *Psychological Science*, 9, 139-142.
- National Center for Education Statistics. (1994). Second follow-up: Student component data file user's manual (NCES 94-374). Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (1996). National educational longitudinal study: 1988-94 data files and electronic codebook system. Base year through third year follow-up ECB/CD-ROM 1996. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (1998). Digest of educational statistics (NCES 1999-032). Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (1999a). The conditions of education 1999. (NCES, 1999-022). Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (1999b). Dropout rates in the United States: 1998. Statistical Analysis Report (NCES 2000-022). Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (1999c). Trends in academic progress: Three decades of student performance. (NCES 2000-469) Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (2000). Early childhood longitudinal study-kindergarten class of 1998-99, Fall 1998. (Statistical Analysis Report 2000, NCES, 2000-070). Washington, DC: U.S. Department of Education.
- Nyberg, K. L., McMillin, J. D., O'Neill-Rood, N., & Florence, J. N. (1997). Ethnic differences in academic retracking: A four-year longitudinal study. *Journal of Educational Research*, 91, 33-41.
- Osbourne, J. W. (1995). Academics, self-esteem, and race: A look at the underlying assumptions of the disidentification hypothesis. *Personality and Social Psychology Bulletin*, 21, 449-455.
- Osbourne, J. W. (1997). Race and academic disidentification. *Journal of Educational Psychology*, 89, 728-735.
- Pergamit, M. R. (1995). Assessing school to work transitions in the United States (U.S. Department of Labor Report NLS 96-32). Washington, DC: U.S. Department of Labor.
- Phillips, M., Crouse, J., & Ralph, J. (1998). Does the Black-White test score gap widen after children enter school? In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 229-272). Washington, DC: Brookings Institute Press.
- Ralph, J., & Crouse, J. (1997). Reading and mathematics achievement: Growth in high school (NCES 98-098). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

- Reynolds, A. J., & Temple, J. A. (1998). Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago Longitudinal Study. *Child Development*, 69, 231-246.
- Rumberger, R. W., & Willms, J. D. (1992). The impact of racial and ethnic segregation on the achievement gap in California high schools. *Educational Evaluation and Policy Analysis*, 14, 377-396.
- Schultz, E. M., & Nicewander, W. A. (1997). Grade equivalent and IRT representations of growth. *Journal of Educational Measurement*, 34, 315-331.
- Sewell, W. H., Haller, A. O., & Portes, A. (1969). The educational and early occupational attainment process. *American Sociological Review*, 34, 82-92.
- Shaycoft, M. F. (1967). Project TALENT Interim Report 3. The high school years: Growth in cognitive skills. Pittsburgh, PA: American Institutes for Research.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613-629.
- Zahs, D., Pedlow, S., Morrissey, M., Marnell, P., & Nichols, B. (1995). High school and beyond second follow-up methodological report. Washington, DC: National Center for Education Statistics, U.S. Department of Education.