RMLE Online
Research in Middle Level Education

David Hough, Ph.D., Editor
Missouri State University
Springfield, Missouri

2004 • Volume 28 • Number 1 ISSN 1084-8959

The Link Between Instructional Practice and the Racial Gap in Middle Schools

Harold Wenglinsky
Hunter College
New York, NY

Citation: Wenglinsky, H. (2004). The link between instructional practice and the racial gap in middle schools. *Research in Middle Level Education Online, 28*(1). Retrieved [date], from http://www.nmsa.org/Publications/RMLEOnline/tabid/101/Default.aspx

Abstract

The No Child Left Behind Act of 2001 calls for improving test scores of minority students in third through eighth grades, making middle level schools a key venue for closing the racial gap in achievement. This study, by analyzing data on the 15,000 eighth graders who took the 2000 National Assessment of Educational Progress (NAEP) in mathematics, provides evidence that middle level schools can make a difference in Black-White and Latino-White test score gaps. Using multi-level Hierarchical Linear Modeling techniques, this study finds that depending upon the instructional practices that middle school teachers utilize, the achievement gap can be reduced substantially for African American and Latino students.

Introduction

No Child Left Behind (NCLB) has put the spotlight as never before on two issues in educational policy: middle school reform and the racial achievement gap. NCLB calls for states to develop systems of accountability in which students in third through eighth grades will be assessed in mathematics and reading; states and schools will be expected to show not just that the student body as a whole, but each racial and ethnic group, is making satisfactory progress. Because the testing occurs primarily during the middle school years, the burden of raising achievement will be felt chiefly by middle schools. If all students are to be proficient by 2013, as NCLB demands, middle school education will need to be reformed. Also, because the legislation specifically calls for the progress of each racial and ethnic group, the ultimate goal is to close the achievement gap. Since the average performance of minority students is lower than that of white students, middle schools will have further to go in moving minority students to proficiency.

States have already begun to develop plans and pass legislation to encourage middle schools to close the achievement gap. A good example is North Carolina, which has begun implementing an 11-point strategy, based upon its report *Implementation Plan for Recommendations from the North Carolina Advisory Commission on Raising Achievement and Closing Gaps* (North Carolina Advisory Commission on Raising Achievement and Closing Gaps, 2003). To close the achievement gap, the strategy includes reducing the disproportionate number of minorities in special education, exposing minority students who are achieving near grade level to advanced and challenging content, providing teachers with professional development on addressing the needs of an ethnically diverse population, improving teacher education to increase the responsiveness of prospective teachers to minority students, providing monetary incentives for those who want to teach in high-need schools, and addressing the achievement gap as part of the accountability system with the goal of having 95% of ethnic minority students reach grade level by 2010. Kentucky, to cite another example, has also recently enacted measures to close the achievement gap, including the creation of biennial targets and the development of school plans using state professional development funds (Christie, 2002).

The current study seeks to investigate ways in which instructional practices of teachers can be modified to reduce the gap. Most prior research on the gap has focused on what can be done for students before they start school and on the mixing of students from schools with high and low minority populations. This study uses data on the 15,000 eighth graders who took the 2000 National Assessment of Educational Progress in mathematics to relate 14 instructional practices to the achievement gap, taking into account student socio-economic status. Using the statistical technique of Hierarchical Linear Modeling, the study finds that some instructional practices can indeed help close the gap, suggesting that middle schools can make a difference by reforming what goes on in the classroom.

Background

Researchers have identified a variety of factors in the achievement gap. These include the situation to which children are exposed before schooling begins, the gap due to demographics that may create a gap in the social dynamics of schools, and the gap attributable to school policies and practices *per se*. Many researchers have suggested that test score gaps are rooted in children's experiences before entering school. Jencks and Phillips (1998) have argued that family experiences and preschool are the key to creating (or limiting) the achievement gap, and they point to a gap that is already significant when students enter kindergarten (also Lee & Burkham, 2002; Phillips, Brooks-Gunn, Duncan, Klebanov & Crane, 1998). The view is that test score gaps are a function of the demographics of students' peers is rooted in the literature on desegregation, which suggests that minority students perform better when they have a significant set of middle class white peers (Moses & Orfield, 2002). But it is also possible that the achievement gap is a function of the policies and practices of individual schools, particularly what occurs between teachers and students in classrooms.

Teachers employ a variety of possible instructional practices. Depending upon which they select, students will perform better or worse on assessments of knowledge. It is only this third view that gives schools the power to close the achievement gap. Social inequalities that exist prior to children's entering school can be addressed in one of two ways: through improving the home environment or providing high quality day care, neither of which is the responsibility of the school system. Social inequalities that create demographic inequalities between schools can also only be remedied through actions outside the purview of the schools[1]. But if part of the problem is the nature of instruction in middle schools, middle school principals can make a difference. By targeting instructional practices that raise the average achievement of the student body, they can improve overall school quality. And by targeting instructional practices that disproportionately benefit minority students, they can help remedy the achievement gap.

Research on instructional practices, however, provides little guidance as to which practices may most profitably be encouraged. Until the mid-1990s, most research on instructional practice was small scale, studying one or a few schools (e.g. Cohen, McLaughlin & Talbert, 1993). The reason was that it was difficult to capture what occurs in the classroom using questionnaires and other instruments used in large-scale research. Not surprisingly, large-scale research limited itself to studying aspects of teaching that were easily measurable, namely the background characteristics of teachers such as their levels of educational attainment and years of experience. The findings of such studies (known as "production functions") regarding the impact of teacher characteristics on student performance, were extremely mixed. Meta-analyses, which summarized the results from hundreds of these studies, themselves came to divergent conclusions (see Hanushek, 1997, 1996a, 1996b, 1989; Hedges, Laine & Greenwald, 1994; Greenwald, Hedges & Laine, 1996; Hedges & Greenwald, 1996 for divergent reviews of the literature.). The two exceptions to this rule were studies of teachers' college majors and teachers' academic proficiency as measured by standardized tests. These two characteristics proved to be strongly associated with student performance (Ferguson, 1991; Ferguson & Ladd, 1996; Monk, 1994; Goldhaber & Brewer, 1996).

In the last decade, however, the emergence of more comprehensive databases has led to large-scale analyses of the impact of instructional practices on student performance. In 1996, the National Educational Longitudinal Study, a nationally representative database, was used to relate a few teacher practices in math and science to student performance in those subjects (National Center for Education Statistics, 1996). It found no relationships in math, but in science it found that students performed better when teachers emphasized higher-order thinking skills. A study by Cohen and Hill (2000) related classroom practices to student performance for the entire state of California and found a link between teachers' emphasizing higher-order thinking skills and student mathematics performance. Using the 1996 NAEP in mathematics, Wenglinsky (2002) found a series of classroom practices, including an emphasis on higher-order thinking skills and hands-on learning, to be positively related to student mathematics performance. Also, Wenglinsky (2003) used the 2000 NAEP in reading and found a link between teaching metacognitive skills and student reading performance.

While large-scale research has linked classroom practices to average student performance it has not found links to the achievement gap. This is due partly to the fact that it must be understood that there are two achievement gaps: the one between schools and the one within schools. The between-school achievement gap stems from the segregated nature of schools; some are predominantly white and some are predominantly minority, with the white schools tending to outperform the minority ones. While much of this gap may be attributable to demographic factors, some may be due to school factors such as instructional practices. Perhaps the culture of a typical predominantly white school is conducive to teachers engaging in a lot of group preparation time and strong mentoring relationships between new and veteran teachers. Such a culture might lead teachers at that school to employ uniformly more effective instructional practice than teachers at a typical minority school with a less collegial faculty. Thus differences in instructional practice between schools might lead to differences in achievement between schools, causing a between-school racial achievement gap. The within-school achievement gap stems from the fact that educational experiences differ both between classes in the same school and between students in the same class. Curricular policies such as tracking may cause students to have different experiences in the same grade in the same school. Variations in teacher quality may have an effect, with the stronger teachers being assigned to more advanced classes and stronger students. And within a classroom, a teacher may be more effective with some kinds of students than others. These differences can be racially based. Some research suggests that more affluent parents are better able to get their children into classes with stronger teachers, and within those classes to get greater attention for their children. And tracking policies often overlap with race. Low-track classes have very often been found to be disproportionately minority.

Only two recent large-scale studies shed light on the interrelationships among instructional practice and racial achievement gaps. One, by Sarah Theule Lubienski (2002) analyzes the National Assessment of Educational Progress in mathematics for fourth, eighth, and twelfth graders in 1990, 1996, and 2000 and quantifies substantial gaps between white and black students, taking student socioeconomic status (SES) into account. The article argues the superiority of this approach to simply comparing black and white students, because it makes it possible to compare blacks and whites on a purely racial dimension, with similar levels of SES. The study does not relate instructional practices to the racial achievement gap, but it does document that most of the instructional practices reformers have identified with high achievement in mathematics are less likely to be used by teachers of black students than by those of white students. The other study (Von Secker, 2002), using the National Educational Longitudinal Study of 1988, did link instructional practices in biology to the racial gap in science scores. The study analyzed 4,377 tenth graders who were taking biology in 1,406 classes using HLM. Five inquiry-based teaching practices were related to the within-class achievement gap between white and minority students. The study found that there was a racial achievement gap associated with many of these practices, and because the practices were inquiry-based in their content, the study concluded that high schools could reduce the racial achievement gap by adopting such practices.

While the latter study constitutes a good first step in research linking instructional practices to the achievement gap, methodological issues limit its usefulness for middle school administrators seeking to close the gap. First, the study is of high school biology. It may be that practices that are developmentally appropriate for high school students are not for middle school students, and biology results may not obtain in the two subjects emphasized by NCLB, math and reading. Second, the range of instructional practices studied, five, was not sufficiently comprehensive; the literature linking instructional practices to average achievement typically uses ten or more. As explained by Mayer (1999), using low numbers of practices makes the measures potentially unreliable and invalid. A third problem is that the study did not control for SES at the school level. Without doing so, there is always the potential that the racial gap is an economic gap, as the Lubienski study points out. Also, it is possible that "effective" instructional practices are really a proxy for high SES students who achieve at a high level, rather than the practices themselves being responsible for high achievement. And finally, the study did not distinguish between the Black-White racial gap and the Latino-White racial gap. It may be that what constitutes effective practice varies not only between whites and minorities but among minorities. The greater likelihood that Latino students are English Language Learners, for instance, might have pedagogical ramifications for how best to close the gap.

How the Current Study Was Conducted

The current study seeks to address these problems to answer two questions pertinent to the middle school racial achievement gap:

1. Do instructional practices affect the achievement gap primarily at the between-school or at the within-school level?
2. What kinds of instructional practice are most effective for reducing the achievement gap?

To answer these questions, this study makes use of data on the 15,694 eighth graders who took the National Assessment of Educational Progress (NAEP) in 2000 in mathematics. NAEP is administered every year in a variety of subjects including math, science, reading, and civics to nationally representative samples of fourth, eighth, and twelfth graders. Referred to as "the Nation's Report Card," NAEP is used to measure how much students know, compare knowledge among subgroups, and follow knowledge over time. In addition to taking an assessment, students fill out a questionnaire, as do teachers and school administrators. The teacher questionnaire includes information on teacher background and classroom practices and the student questionnaire includes student demographic information. It is therefore possible to combine information about student test scores, student SES, student race, teacher background, and instructional practices to relate the practices to the two types of achievement gap, between- and within-school.

The present study analyzes these data using the technique of Hierarchical Linear Modeling (HLM). The basic principle behind HLM is that any given student characteristic being analyzed exists at two levels of aggregation: the student and the school (Bryk & Raudenbush, 1992). For instance, the SES of an individual student may have an effect on his or her test scores, a student-level effect, and the SES of his or her peers at the school may have an effect, a school-level effect[2]. HLM estimates three sets of equations:

- Student level demographics are related to individual student test scores.
- Average school test scores are related to school aggregates of teacher and student characteristics.
- Each of the relationships between student-level demographics and student test scores (their "slopes") are themselves related to the school aggregates, with one equation for each slope[3].

In this study, two HLMs are developed, one to estimate the racial achievement gap and one to estimate the impact of instructional practices on that gap.

First HLM: For this model, the first equation is the student-level equation relating individual student test scores to ($\boldsymbol{y}_{0j}$) school average test scores (β_{0j}), a student being African American (β_{1j}), a student being Latino (β_{2j}), student SES (β_{3j}), and an error term at the student level ($\boldsymbol{r}_{ij}$). The second equation is the first school-level equation, relating school average test scores from the first equation (β_{0j}) to the intercept (γ_{00}), the percentage of students in the school who are African American (γ_{01}), the percentage of students in the school who are Latino (γ_{02}), the average SES (γ_{03}) and the school level error (u_0). The third, fourth, and fifth equations merely relate the slopes from the first equation to the corresponding school error terms (u_n). This specification distinguishes between the within-school racial achievement gaps (β_{1j}, β_{2j}) and the between-school racial achievement gaps (γ_{01}, γ_{02}); between African American-White and Latino-White gaps (β_{1j}, γ_{01}) and (β_{2j}, γ_{02}).

The gaps are net of SES, indicating that they are racial differences in achievement for students or schools with similar SES (SES is included in the equations). The numbers refer to points on the NAEP scale, where 12 points is roughly one grade level.

Second HLM: For this model, the first equation remains the same as in the first model. To the second equation, off of β_{0j} is added slopes for the instructional practices ($\gamma_{11} \ldots \gamma_{1n}$; $\gamma_{21} \ldots \gamma_{2n}$). The fifth equation remains unchanged from the first HLM. The total impact of instructional practice on between-school racial gaps is the difference in γ_{01} for African Americans and γ_{02} for Latinos between the two models. The total impact of instructional practice on within-school racial gaps is the difference in β_{1j} for African Americans and β_{2j} for Latinos between the two models. The impact of each particular instructional practice on between-school racial gaps is its coefficient in the β_{0j} equation, or γ_{0n}. The impact of each particular instructional practice on within-school racial gaps is its coefficient in the corresponding β_{mj} equation, or γ_{mn} [4].

This approach addresses the problems of the HLM regarding the racial gap in the high school biology study. It includes separate estimates for within- and between-school gaps; it distinguishes between African American and Latino gaps; it includes a large number of instructional practices (as will be seen); and the racial achievement gaps are net of SES.

Results

Before discussing the results of the HLMs, it is worthwhile to examine the relative frequency of the various instructional practices under study (see Table 1 for variable definitions and descriptive statistics). In total, 14 practices pertaining to mathematics instruction were included. These were: two items about time spent on math, class time, and homework time; four items on the conceptual emphasis of the teacher, emphasis on facts, emphasis on rote learning, emphasis on mathematical reasoning, and emphasis on communicating math ideas; and eight teaching techniques, using textbooks, engaging in group work, working with objects, taking tests, writing about math, talking about math, doing math projects, and solving problems grounded in real world situations. Time spent on math by eighth graders is relatively high, averaging more than 2.5 hours per day in class and a half hour per day at home. This amount of time does not seem to vary much from school to school, student to student (standard deviations are well under 1). In terms of emphasis, there seems to be greater emphasis on basic skills approaches (facts, rote) than on higher-order thinking (reason, communication). With regard to the techniques, the most common is using textbooks, followed by real world problem solving, talking about math, and working in groups. The relatively uncommon techniques are working on projects and writing about math.

In addition to the instructional practices, the study examined teacher and student background characteristics (Table 2). The typical teacher had more than six years of experience, although experience varied greatly (SD=3.31). Four out of ten teachers had a masters degree or higher. And just 18% had majored in mathematics or math education. Among students, 15% were African American and 15% Latino. Their SES was measured based upon the mothers' and fathers' levels of education, as well as whether there were various reading materials in the home (newspapers, magazines, encyclopedias, and books). The typical parent had received some college education, but not enough to graduate and at least three out of four households had each of the reading materials.

The first HLM reveals the magnitude of the achievement gap and partitions it into four components: between-school African American, between-school Latino, within-school African American and within-school Latino (Table 3). The school mean NAEP score is 209 points. Schools with higher percentages of African American students have school means that are 38 points lower. Schools with higher percentages of Latino students have school means that are 17 points lower. Within schools, African American students score 25 points lower than their white peers and Latino students score 15 points lower. It is possible to total the Latino and African American effects because students who identify themselves as African American are also identifying themselves as not Latino. Thus the school level racial achievement gap is 54 points and the student level gap is 41 points. Thus, the gap seems greatest between predominantly African American and predominantly white schools, but is substantial and statistically significant for all groups, even taking school and student SES into account.

Adding the instructional practices and teacher background to the school-level intercept model (β_{0j}) indicates that some instructional practices are related to school-level achievement (Table 4, column 2). The more time students spend on homework, the better the school performs. Also, the more students do real-world problems, the better the school performs. A couple of techniques, however, depress scores. First, the more testing that occurs in the school, the worse students perform on NAEP. Second, writing about math seems to have a negative effect on the intercept. The inclusion of the various instructional practices also does not explain a substantial portion of the school-level achievement gap. Compared to the first HLM, the second actually has a slightly higher between-school black-white test score gap, and an equivalent between school Latino-white test score gap. Thus instructional practice cannot be said to reduce the racial achievement gap between schools.

Adding the instructional practices and teacher background to the student level achievement slopes for African Americans (β_{1j}) and Latinos (β_{2j}), however, does show substantial reduction in the slopes, and therefore the gaps. (A negative number indicates a reduction in the gap). For African Americans (column 3) time-on-task (in math) in class reduces the achievement gap by four points. One teacher characteristic, having a masters degree, actually raises the gap by three points, suggesting that better educated teachers may be less responsive to the needs of low-achieving students. For Latinos, working on projects reduces the achievement gap by nearly six points. All totaled, instructional practices reduce the African American within-school gap by six points (25-19) and the Latino within-school gap by 11 points (15-4). Indeed, the within-school Latino gap is down to just over four points. It should also be noted that the within school gaps for both groups are sufficiently small to be statistically insignficant.

Conclusions

Before interpreting these findings, it is important to note shortcomings of the present study. First, the data are cross-sectional. This means that nothing is known about the causal direction of the results. While the research questions assume that instructional practices are having an effect on certain racial gaps, it is possible that teachers chose certain instructional practices in response to the way achievement is distributed across racial groups of students. For instance, it is possible that teachers choose to spend more time on mathematics in classes where math achievement is more racially homogenous. Second, while the study represents a substantial number of instructional practices, it in no way replicates the detail or nuance of classroom observations. Subsequent research should attempt to code such observations on a large scale and relate them to the achievement gap, on the model of the work done with video on the teaching gap.

These caveats notwithstanding, the study makes significant contributions to the understanding of the racial achievement gap in middle schools. First, it demonstrates that what occurs prior to middle school is not graven in stone. The study gauges the impact of eighth grade mathematics teachers only, and finds that in this single year the racial divide within a school can be reduced. Thus, the view that minority students are so far behind by the end of preschool, let alone by the end of elementary school, that schools can do nothing to help them catch up is belied. Depending upon what techniques they employ, teachers can close the gap, at least within their schools. That is the second finding. Many have viewed the key action for reducing racial inequality as changing the demographic composition of the schools through desegregation. African American and Latino students, it is argued, will benefit by being exposed to their white peers. Whether this is true or not, a revival of desegregation policies seems unlikely. "White flight" from urban areas means that desegregation must involve interdistrict busing. And the move to choice in urban areas has led to greater segregation, as minority students flee to certain schools. Thus, the population of many cities, and of many schools within those cities, is overwhelmingly minority. That said, this study suggests that, even in the context of a growing urban minority population, urban schools that have some ethnic diversity do have the power to reduce the gap by eliminating differences between minority and white students within their schools.

References

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park: Sage Publications.

Christie, K. (2002). States address achievement gaps. *Phi Delta Kappan, 84*(2); 102-103.

Cohen, D. K., & Hill, H. C. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record, 102*(2), 294-343.

Cohen, D. K., McLaughlin, M. W., & Talbert, J. E. (1993). *Teaching for understanding: Challenges for policy and practice.* San Francisco: Jossey-Bass.

Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. In H. F. Ladd (Ed.), *Holding school accountable: Performance-based reform in education* (pp. 265-298). Washington, DC: The Brookings Institution.

Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal of Legislation, 28*(2), 465-498.

Goldhaber, D. D., & Brewer, D. J. (1996). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of Human Resources, 32*(3), 505-520.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361-396.

Hanushek, E. A. (1989). The impact of differential expenditures on school performance. *Educational Research, 18*(4), 45-51.

Hanushek, E. A. (1996a). A more complete picture of school resource policies. *Review of Educational Research, 66*(3), 397-409.

Hanushek, E. A. (1996b). School resources and student performance. In G. T. Burtless (Ed.), *Does money matter? The effect of school resources on student achievement and adult success* (pp. 43-73). Washington, DC: The Brookings Institution.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19*(2), 141-164.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Research, 23*(3), 5-14.

Hedges, L. W., & Greenwald, R. (1996). Have times changed? The relation between school resources and student performance. In G. T. Burtless (Ed.), *Does money matter? The effect of school resources on student achievement and adult success* (pp. 74-92). Washington, DC: The Brookings Institution.

Jencks, C., & Phillips, M. (1998). Introduction. In C. Jencks & M. Phillips (Eds.) *The Black-White test score gap* (p. 1-51). Washington, DC: Brookings Institution Press.

Johnson, E., Mislevy, R. J., & Thomas, N. (1994). Scaling procedures. In E. Johnson & J. Carlson, *The NAEP 1992 technical report* (pp. 241-256). Princeton, NJ: Educational Testing Service.

Johnson, E. G. (1989). Considerations and techniques for the analysis of NAEP data. *Journal of Education Statistics, 14*(4), 303-334.

Lee, V. E., & Burkham, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school.* Washington, DC: Economic Policy Institute.

Lubienski, S. T. (2002). A closer look at Black-White mathematics gaps: Intersections of race and SES in NAEP achievement and instructional practices data. *Journal of Negro Education, 71*(4), 269-287.

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis, 21*(1), 29-45.

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13*(2), 125-145.

Moses, M., & Orfield, G. (2002). *Embracing race: Why we need race-conscious education policy*. New York: Teachers College Press.

National Center for Education Statistics. (1996). *High school seniors' instructional experiences in science and mathematics*. Washington, DC: U.S. Government Printing Office.

North Carolina Advisory Commission on Raising Achievement and Closing Gaps. (2003). *Implementation plan for recommendations from the North Carolina Advisory Commission on Raising Achievement and Closing Gaps*. www.ncpublicschools.org/schoolimprovement/closingthegap/reports/downloads/implementation03.pdf

O'Reilly, P. E., Zelenak, C. A., Rogers, A. M., & Kline, D. L. (1996). *1994 trial state assessment program in reading secondary-use data files user guide*. Washington, DC: U.S. Department of Education.

Phillips, M., Brooks-Gunn, J., Duncan, G. J., Klebanov, P., & Crane, J. (1998). Family background, parenting practices and the Black-White test score gap. C. Jencks & M. Phillips (Eds.) *The Black-White test score gap* (pp. 103-145). Washington, DC: Brookings Institution Press.

Von Secker, C. (2002). Effects of inquiry-based teacher practices on science excellence and equity. *Journal of Educational Research, 95*(3), 151-160.

Wenglinsky, H. (2003). Using large-scale research to gauge the impact of instructional practices on student reading comprehension: An exploratory study. *Education Policy Analysis Archives, 11,* 19.

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives, 10,* 12.

Endnotes

[1]To the extent that segregated schools are inherently unequal, only the intervention of the courts, through busing programs and the like, can increase racial heterogeneity.

[2]For purposes of this paper an "effect" does not assume a particular causal direction for a relationship, but merely the existence of such a relationship.

[3]In the HLM, the first equation relates student level variables (test scores and student background) to one another, with student background varying from an intercept, as follows:

$$(1)\ \boldsymbol{y}_{ij} = \beta_{0j} + \beta_{1j}\boldsymbol{x}_{ij} + \boldsymbol{r}_{ij}$$

where

$\boldsymbol{y}_{ij}$ is the student-level variation in test scores

β_{0j} is the intercept, or the mean test score for a school

β_{1j} is the relationship between student-level variation in student background and student-level variation in test scores

$\boldsymbol{x}_{ij}$ is student-level variation in student background and

$\boldsymbol{r}_{ij}$ is student-level variation other than student-level variation in student background.

The second equation relates the school or classroom level independent variables (teacher and student background and classroom practices) to school or classroom level variation in test scores. School-level variation in test scores is represented by β_{0j} because it consists of variation in test scores absent the student-level variation separated out as $\beta_{0j}\boldsymbol{x}_{ij}$ and $\boldsymbol{r}_{ij}$. The second equation is thus:

$$(2)\ \beta_{0j} = \boldsymbol{y}_{00} + \boldsymbol{y}_{01}\boldsymbol{w}_{1} + \ldots + \boldsymbol{y}_{0n}\boldsymbol{w}_{n} + \boldsymbol{u}_{0}$$

where

β_{0j} is as in equation (1)

$\boldsymbol{y}_{00}$ is the intercept, or the grand mean test scores absent variation by school

$\boldsymbol{y}_{01}$ is the relationship between school-level variation in a teacher or classroom characteristic and school-level variation in test scores

$\boldsymbol{w}_{1}$ is a classroom or teacher characteristic

$\boldsymbol{y}_{0n}$ is the relationship of the nth classroom or teacher characteristic to school-level variation in test scores

$\boldsymbol{w}_{n}$ is the nth classroom or teacher characteristic and

$\boldsymbol{u}_{0}$ is the school-level variation in achievement unexplained by the n coefficients.

The third equation relates the school or classroom level independent variables (teacher and student background and classroom practices) to the relationship between test scores and student background. Using this relationship as a dependent variable makes it possible to gauge the impact of instructional practice on the polarization of achievement above and beyond average school achievement. This relationship is represented by as per equation (1). The third equation is thus:

(3) $\beta_{1j}=\boldsymbol{y}_{10}+\boldsymbol{y}_{11}\boldsymbol{w}_{1}+...+\boldsymbol{y}_{1n}\boldsymbol{w}_{n}+\boldsymbol{u}_{1}$

where

β_{1j} is as in equation (1)

$\boldsymbol{y}_{10}$ is the intercept, or the grand mean test scores absent variation by school

$\boldsymbol{y}_{11}$ is the relationship between school-level variation in a teacher or classroom characteristic and school-level variation in test scores

$\boldsymbol{w}_{1}$ is a classroom or teacher characteristic

$\boldsymbol{y}_{1n}$ is the relationship of the nth classroom or teacher characteristic to school-level variation in test scores

$\boldsymbol{w}_{n}$ is the nth classroom or teacher characteristic and

$\boldsymbol{u}_{1}$ is the school-level variation in achievement unexplained by the n coefficients.

The third equation may be four, five, or more equations depending upon the number of student background characteristics included. In this case, where the background characteristics are SES, being African American, or being Latino, there are a total of five equations.

[4]Certain methodological issues arise from the use of NAEP for these analyses. First, NAEP does not provide a single test score for each student. Each student takes only a small subset of the test, and consequently the test score for a particular student needs to be imputed using a procedure known as plausible values methodology. The end result is five test scores rather than one, and separate HLMs have to be run for each test score and combined into a final model. Second, NAEP is not a simple random sample, but, rather, clusters students within schools, which are clustered within primary sampling units, consisting of one or a few school districts. Because of this, HLM and other techniques may underestimate standard errors, treating as statistically significant relationships that are not. Consequently, the standard errors have to be inflated by what is known as a design effect to determine whether the relationships are actually statistically significant (Johnson, 1989; Johnson, Mislevy & Thomas, 1994; O'Reilly, Zelenak, Rogers & Klein, 1996).

Figures and Tables

TABLE 1. Descriptive Statistics for Instructional Practices

Practice	**M**	**SD**	**N**
Time per week on math instruction (1=2.5 hrs or less; 3=4 hrs or more)	2.30	.64	11981
Math homework assigned/day (1=none; 6=more than 1 hr)	2.99	.80	11916
Emphasis on math facts (1=little/no emphasis; 3=heavy emphasis)	2.74	.51	11837
Emphasis on solving rote problems (1=little/no emphasis; 3=heavy emphasis)	2.79	.44	11826
Emphasis on reasoning (1=little/no emphasis; 3=heavy emphasis)	2.54	.57	11806
Emphasis on communicating (1=little/no emphasis; 3=heavy emphasis)	2.34	.65	11791
Use Textbook (1=never/hardly ever; 4=almost every day)	3.73	.59	11841
Working in groups (1=never/hardly ever; 4=almost every day)	2.74	.90	11847
Working with objects (1=never/hardly ever; 4=almost every day)	2.12	.81	11811
Taking math tests (1=never/hardly ever; 4=almost every day)	2.43	.53	11844
Writing about math (1=never/hardly ever; 4=almost every day)	2.05	.92	11831
Talking about math (1=never/hardly ever; 4=almost every day)	2.83	1.15	11862
Working on projects (1=never/hardly ever; 4=almost every day)	1.39	.57	11849
Solving real world problems (1=never/hardly ever; 4=almost every day)	2.96	.83	11867

TABLE 2. Descriptive Statistics for Student and Teacher Background

Characteristic	M	SD	N
Student Socioeconomic Status			
Mother's Education Level (1= < H.S Degree; 4=College Degree)	3.05	1.03	13452
Father's Education Level (1= < H.S. Degree; 4=College Degree)	3.09	1.04	12198
Subscribe to Newspaper (1=yes;0=no)	.74	.44	14987
Own Encyclopedia (1=yes;0=no)	.85	.35	14945
Own 25+ Books (1=yes;0=no)	.95	.22	15012
Teacher Background			
Years of Experience (1=2 or less; 4=25 or more)	3.32	1.32	13293
Education Level (1=Masters or more; 0=Less than Masters)	.41	.49	13306
Teacher Major (1=Math or Math Education; 0=Other)	.18	.14	15694

TABLE 3. Hierarchical Linear Model for Measurement of Racial Achievement Gap

School-level Demographic	Mean School Achievement	Student African American	Student Latino	Student SES	Student Error (SD)
Intercept	208.99** (4.51)	-25.24** (1.08)	-15.58** (1.16)	1.61** (.14)	28.99
% African American	-37.85** (1.99)				
% Latino	-17.67** (2.48)				
Average SES	8.08** (.44)				
School Error (SD)	10.70	9.92	11.40	.39	

*p< .10; **p< .05

Except for residuals, cells contain unstandardized coefficients and standard errors

School-level Demographic	**Mean School Achievement**	**Student African American**	**Student Latino**	**Student SES**	**Student Error (SD)**
Intercept	195.06** (8.23)	-19.31 (18.74)	-4.13 (18.87)	1.61 (.13)	28.99
% African American	-38.56** (2.04)				
% Latino	-17.77** (2.50)				
Average SES	7.59** (.46)				
Teacher Experience	.48 (.50)	-.30 (1.23)	.28 (1.19)		
Teacher Degree	.25 (1.30)	9.10** (3.33)	-.12 (3.34)		
Teacher Major	-.11 (4.95)	5.93 (12.76)	10.13 (13.40)		
Time in Class on Math	-1.00 (.91)	-4.65** (-2.17)	-1.85 (2.43)		
Time on Homework	2.57** (.85)	-.99 (2.04)	.67 (2.12)		
Textbook	1.69 (1.11)	-2.30 (2.47)	-2.25 (2.28)		
Work in Groups	.43 (.71)	.28 (1.73)	-1.10 (1.68)		
Work with Objects	-.34 (.82)	-1.64 (2.12)	1.53 (1.97)		
Take Tests	-2.07* (1.22)	3.93 (2.99)	-.32 (2.77)		
Write about Math	-1.56 (.80)	.44 (1.86)	-1.11 (1.90)		
Talk about Math	-.07 (.58)	-1.30 (1.47)	-1.14 (1.53)		
Do Math Projects	.38 (1.25)	1.72 (2.87)	-5.78* (3.06)		
Solve Real World Problems	2.56** (.90)	-.64 (2.35)	.89 (2.04)		
Emphasize Facts	.49 (1.45)	.63 (3.54)	2.84 (3.57)		
Emphasize Rote	-1.40 (1.75)	2.52 (4.88)	-4.32 (4.48)		
Emphasize Reasoning	1.46 (1.32)	-3.19 (3.17)	2.87 (3.12)		
Emphasize Communication	1.72 (1.15)	3.22 (2.88)	2.70 (2.84)		
School Error (SD)	10.47	8.83	11.64		

*p< .10; **p< .05

Except for residuals, cells contain unstandardized coefficients and standard errors