EDUCATION POLICY ANALYSIS ARCHIVES

A peer-reviewed scholarly journal
Editor: Gene V Glass
College of Education
Arizona State University

Volume 12 Number 64 November 23, 2004 ISSN 1068-2341

Closing the Racial Achievement Gap: The Role of Reforming Instructional Practices

Harold Wenglinsky
Hunter College

Citation: Wenglinsky, H. (2004, November 23) Closing the racial achievement gap: The role of reforming instructional practices. *Education Policy Analysis Archives, 12*(64). Retrieved [date] from http://epaa.asu.edu/epaa/v12n64/.

Abstract

No Child Left Behind calls for schools to close the achievement gap between races in math and reading. One possible way for schools to do so is to encourage their teachers to engage in practices that disproportionately benefit their minority students. The current study applies the technique of Hierarchical Linear Modeling to a nationally representative sample of 13,000 fourth graders who took the 2000 National Assessment of Educational Progress in mathematics to identify instructional practices that reduce the achievement gap. It finds that, even when taking student background into account, various instructional practices can make a substantial difference.

Introduction

The Elementary and Secondary Education Act Reauthorization of 2001, popularly known as "No Child Left Behind (NCLB)," has put the spotlight as never before on the issue of the racial achievement gap. NCLB calls for schools to make "Adequate Yearly Progress (AYP)," meaning that their students' scores on standardized tests are expected to improve from year to year. But not only is the average score of the student body expected to improve; so too are the scores of various demographic subgroups in the student body, including racial and ethnic minorities. And, by 2013, all subgroups of student are expected to be "proficient," meaning that the gap must be eliminated by that time (Education Commission of the States, 2003; Olson, 2003).

Closing the racial achievement gap would be no small feat, given its long history. Data have been available since 1969, when the first National Assessment of Educational Progress (NAEP) was administered in science to nationally representative groups of 9, 13 and 17 year olds (National Center for Education Statistics, 2000). Also known as "the Nation's Report Card," NAEP has since been administered every year or two in a variety of subjects; math began in 1970 and reading in 1971. In addition to measuring aggregate trends in student performance, NAEP measures differences among demographic subgroups. With large initial racial gaps in the early 1970s, NAEP documented their reduction over the 1970s and 1980s. In the 1990s, however, they began to increase again (Lee, 2002). To close the achievement gap by 2013 would involve reversing this trend and dropping score differences to zero in every state in less than a decade, a daunting task.

States are already beginning to pursue a variety of strategies to reduce the gap. A good example is North Carolina, which has begun implementing an eleven-point strategy, based upon its report *Implementation Plan for Recommendations from the North Carolina Advisory Commission on Raising Achievement and Closing Gaps.* 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).

Rather than pursue such a diverse set of strategies, some policymakers and educators have called for focusing on the activities of schools that have the most direct impact on their students: the classroom practices of teachers. As Kati Haycock has argued (2003; see also Bell, 2003), teachers can choose from a variety of strategies to enhance student learning. Yet some of these strategies are presumably more effective than others. And, presumably, there are also particular strategies that will benefit minority students, thus reducing the achievement gap. The problem is that while large-scale research has succeeded in identifying some instructional practices that are beneficial for the student body as a whole, it has generally not succeeded at identifying practices that provide disproportionate benefits to African American and Latino students.

The current study seeks to address this gap in the literature. It makes use of data on the 13,000 fourth graders who took the NAEP math assessment in 2000. Using NAEP has the advantages not only of being the only relatively recent national sample with data on student achievement, student background and instructional practice, but that NCLB intends to use NAEP as the ultimate yardstick for state assessments of AYP. Under NCLB, states will be allowed to use their own assessments, but their results need to be consistent with NAEP. Thus, by identifying instructional practices that are associated with high performance on NAEP, this study can help educational administrators in identifying practices that can be expected to raise scores on their state tests as well. Using the technique of Hierarchical Linear Modeling (HLM), this study finds, as Haycock (2003) anticipated, that a series of instructional practices, when used in concert, can substantially reduce both the black-white and Latino-white achievement gaps.

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 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 et al., 1998). The view 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, 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 which 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 school.[1] But if part of the problem is the nature of instruction in schools, 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).

---

[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.

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 more effectively reach 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 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 school administrators seeking to close the gap. First, the study is of high school biology. It may be that practices which are developmentally appropriate for high school students are not appropriate for younger students, particularly those to be tested under NCLB (third through eighth graders). Also, 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 in order to answer two questions pertinent to the 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 13,511 fourth 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 (see National Center for Education Statistics, 2003). 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.[2]

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.[3] HLM estimates three sets of equations:

1. Student level demographics are related to individual student test scores.
2. Average school test scores are related to school aggregates of teacher and student characteristics.
3. Each of the relationships between student-level demographics and student test scores (their "slopes") is itself related to the school aggregates, with one equation for each slope.[4]

---

[2] It should be noted that teacher self-reports of their practices are often inaccurate, due to teachers either thinking they are engaging in practices in which they are not engaging or because they respond in the way they think the researcher wants. Nonetheless, research has found that teacher self-reports are highly correlated, if not perfectly correlated with classroom practices in mathematics (Mayer, 1999).

[3] 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.

[4] 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) $Y_{ij} = b_{0j} + b_{1j} X_{ij} + r_{ij}$, where

Y_{ij} is the student-level variation in test scores

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

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

X_{ij} is student-level variation in student background, and

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 b_{0j} because it consists of variation in test scores absent the student-level variation separated out as $b_{1j} X_{ij}$ and r_{ij}. The second equation is thus:

(2) $b_{0j} = g_{00} + g_{01} W_1 + ... + g_{0n} W_n + u_0$, where

b_{0j} is as in equation (1)

g_{00} is the intercept, or the grand mean test scores absent variation by school

g_{01} is the relationship between school-level variation in a teacher or classroom characteristic and school-level variation in test scores

W_1 is a classroom or teacher characteristic

g_{0n} is the relationship of the nth classroom or teacher characteristic to school-level variation in test scores

W_n is the nth classroom or teacher characteristic, and

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 $b_{1j} X_{ij}$ as per equation (1). The third equation is thus:

(3) $b_{1j} = g_{10} + g_{11} W_1 + ... + g_{1n} W_n + u_1$, where

b_{1j} is as in equation (1)

g_{10} is the intercept, or the grand mean test scores absent variation by school

g_{11} is the relationship between school-level variation in a teacher or classroom characteristic and school-level variation in test scores

W_1 is a classroom or teacher characteristic

g_{1n} is the relationship of the nth classroom or teacher characteristic to school-level variation in test scores

W_n is the nth classroom or teacher characteristic, and

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.

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 (Y_{ij}) to school average test scores (b_{0j}), a student being African American (b_{1j}), a student being Latino (b_{2j}), student SES (b_{3j}), and an error term at the student level (r_{ij}). The second equation is the first school-level equation, relating school average test scores from the first equation (b_{0j}) to the intercept (g_{00}), the percentage of students in the school who are African American (g_{01}), the percentage of students in the school who are Latino (g_{02}), the average SES (g_{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 (b_{1j}, b_{2j}) and the between-school racial achievement gaps (g_{01},g_{02}); between African American-white and Latino-white gaps (b_{1j},g_{01}) and (b_{2j},g_{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 b_{0j} is added slopes for the instructional practices ($g_{11} \ldots g_{1n}; g_{21} \ldots g_{2n}$). The fifth equation remains unchanged from the first HLM. The total impact of instructional practice on between-school racial gaps is the difference in g_{01} for African Americans and g_{02} for Latinos between the two models. The total impact of instructional practice on within-school racial gaps is the difference in b_{1j} for African Americans and b_{2j} for Latinos between the two models. The impact of each particular instructional practice on between-school racial gaps is its coefficient in the b_{0j} equation, or g_{0n}. The impact of each particular instructional practice on within-school racial gaps is its coefficient in the corresponding b_{mj} equation, or g_{mn}.[5]

This approach addresses the problems of the HLM of 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.

It should be noted that no normative judgment is being made by this research design regarding whether reducing the racial achievement gap is a good thing. Because the racial gaps are net of SES, one might make the argument that, as a matter of educational equity, it is unjust for certain students to trail other students academically simply by virtue of the race into which they are born, and that changing instructional practices from those that contribute to this situation to those that ameliorate it would be a moral good. That said, this study does not seek to make a normative judgment about the existence of a racial achievement gap. It is sufficient justification for studying such a gap that the issue has received renewed attention in recent years as a result of NCLB and state legislation, and that many policymakers would like to know ways to address it.

---

[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).

Results

Before identifying the practices that proved most effective in closing the racial achievement gap, it is worth examining the prevalence of the various practices measured by NAEP (Table 1). In total, 20 practices were analyzed here, two measures of time spent on math work, four on the philosophy of the teacher regarding student learning, five on the content emphasized and nine on the techniques employed. On average, between 2.5 and 4 hours of class time was spent on math and 15-30 minutes of homework assigned. Of the four teacher beliefs, the two most common were emphasizing facts and having students work on routine exercises. Emphasizing math reasoning and communicating math concepts were much rarer. Of the math topics, the one most heavily emphasized was numbers and operations, while algebra, geometry and data analysis were the least emphasized. Of the techniques, the most popular was having students work from textbooks. The least popular were having students work on projects or do hands-on learning with blocks.

Various background characteristics of students and teachers were also included in the models (Table 2). Student socioeconomic status (SES) was measured from whether the student qualified for free or reduced price school lunches, whether the household subscribed to a newspaper or magazines, and whether there was an encyclopedia or books in the home. The teacher background characteristics were whether the teacher had a master's degree or higher, the number of years of teaching experience, and whether the teacher had majored or minored in mathematics or mathematics education at the undergraduate or graduate level.

The first HLM was designed to measure the between- and within-school racial gaps for African American and Latino students, taking their SES into account (Table 3). The model reveals that the average fourth grader scored 193 points with a 27 point gap for African Americans between-school, a 16 point gap for African Americans within-school, a 16 point gap for Latinos between-school and an 8 point gap for Latinos within-school. In other words, the largest gap is between majority black and majority white schools, and the smallest between Latino and white students within the same school.

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)	2.58	.63	11080
Math homework assigned/day (1=none; 6=more than 1 hr)	2.54	.72	11032
Emphasis on math facts (1=little/no emphasis; 3=heavy emphasis)	2.93	.27	11136
Emphasis on solving routine problems (1=little/no emphasis; 3=heavy emphasis)	2.90	.30	11135
Emphasis on reasoning (1=little/no emphasis; 3=heavy emphasis)	2.55	.54	11121
Emphasis on communicating (1=little/no emphasis; 3=heavy emphasis)	2.37	.63	11094
Emphasis on numbers and operations (1=little/no emphasis; 3=heavy emphasis)	2.88	.33	11117
Emphasis on measurement (1=little/no emphasis; 3=heavy emphasis)	2.22	.49	11089
Emphasis on geometry (1=little/no emphasis; 3=heavy emphasis)	2.09	.53	11060
Emphasis on working with data (1=little/no emphasis; 3=heavy emphasis)	1.93	.63	11019
Emphasis on algebra (1=little/no emphasis; 3=heavy emphasis)	1.87	.65	11116
Use Textbook (1=never/hardly ever; 4=almost every day)	3.59	.80	11118
Working in groups (1=never/hardly ever; 4=almost every day)	2.80	.89	11150
Working with objects (1=never/hardly ever; 4=almost every day)	2.56	.78	11150
Working with blocks (1=never/hardly ever; 4=almost everyday)	2.02	.78	11137
Taking math tests (1=never/hardly ever; 4=almost every day)	2.43	.58	11141
Writing about math (1=never/hardly ever; 4=almost every day)	2.24	.96	11171
Talking about math (1=never/hardly ever; 4=almost every day)	2.96	1.10	11145
Working on projects (1=never/hardly ever; 4=almost every day)	1.32	.56	11112
Solving real world problems (1=never/hardly ever; 4=almost every day)	3.08	.86	11172

Table 2

Descriptive Statistics for Student and Teacher Background

Characteristic	**M**	**SD**	**N**
Student Socioeconomic Status			
Qualifies for free or reduced-price lunch (1=yes;0=no)	.61	.49	8627
Subscribe to Newspaper (1=yes;0=no)	.78	.41	10917
Own Encyclopedia (1=yes;0=no)	.81	.39	11067
Own 25+ Books (1=yes;0=no)	.95	.23	12523
Subscribe to Magazine	.80	.40	11383
Teacher Background			
Years of Experience (1=2 or less; 4=25 or more)	3.30	1.29	11957
Education Level (1=Masters or more; 0=Less than Masters)	.34	.47	11858
Teacher Major (1=Math or Math Education;0=Other	.19	.61	13511

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	193.91** (4.40)	-15.65** (1.24)	-7.96** (1.25)	1.77** (.31)	24.44
% African American	-26.82** (2.62)				
% Latino	-16.23** (3.21)				
Average SES	10.40** (.97)				
School Error (SD)	10.25	10.59	9.22	.76	

*p<.10;**p<.05
Except for residuals, cells contain unstandardized coefficients and standard errors.

Table 4
Hierarchical Linear Model for Instructional Practices and Achievement Gap

School-level Demographic	Mean School Achievement	Student African American	Student Latino	Student SES	Student Error (SD)
Intercept	191.22** (12.63)	-9.02 (25.54)	26.85 (22.33)	1.77** (.31)	24.45
% African American	-26.21** (2.62)				
% Latino	-16.05** (3.19)				
Average SES	10.77** (.97)				
Teacher Experience	.95 (.60)	.18 (1.49)	.33 (1.32)		
Teacher Major	-.55 (1.20)	.43 (2.67)	-.93 (2.87)		
Teacher Degree	-.31 (1.78)	-3.07 (4.28)	-.12 (3.74)		
Time in Class on Math	2.67** (1.26)	-4.01 (3.33)	-3.97 (2.91)		
Time on Homework	.52 (1.09)	3.69 (2.39)	-1.67 (2.28)		
Textbook	.30 (.91)	-1.16 (1.95)	-2.13 (1.88)		
Work in Groups	-.89 (.99)	1.69 (2.47)	2.33 (2.14)		
Work with Objects	.23 (1.28)	-1.61 (3.10)	-3.03 (3.01)		
Work with Blocks	.87 (1.21)	2.37 (2.82)	.97 (2.64)		
Take Tests	-2.77** (1.36)	5.59* (3.15)	-.38 (3.44)		
Write about Math	.50 (.95)	-.94 (2.37)	-2.14 (2.10)		
Talk about Math	1.36 (.86)	2.76 (2.13)	.89 (1.96)		
Do Math Projects	-3.95** (1.63)	-1.13 (3.94)	-.20 (3.38)		
Solve Real World Problems	.50 (1.10)	1.04 (2.42)	-2.41 (2.64)		
Emphasize Facts	-15.16** (4.00)	-5.77 (8.66)	2.78 (8.35)		
Emphasize Routine Problems	6.02* (3.52)	3.61 (8.51)	10.90 (7.90)		
Emphasize Reasoning	.03 (1.81)	2.47 (5.09)	-5.49 (3.99)		
Emphasize Communication	-.54 (1.62)	1.71 (4.47)	-2.76 (4.01)		
Emphasize Numbers	2.56 (2.71)	3.27 (6.52)	-10.25 (6.78)		
Emphasize Measurement	2.46 (1.96)	-8.54** (4.91)	-1.42 (4.49)		
Emphasize Geometry	2.89* (1.70)	-5.02 (4.20)	-2.17 (3.94)		
Emphasize Data	.15 (1.44)	1.20 (3.64)	6.48** (3.26)		
Emphasize Algebra	1.64 (1.35)	.89 (3.31)	1.18 (3.14)		
School Error (SD)	9.65	9.99	8.70	.79	

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

Except for residuals, cells contain unstandardized coefficients and standard errors

The second HLM finds that when instructional practices are introduced, the within-school gaps go away while the between-school gaps remain unchanged. The coefficients for African American and Latino schools (Column 2, Table 4) are not substantially different from those in the first HLM. This indicates that the introduction of instructional variables does not mitigate the advantage of predominantly white schools over predominantly African

American or predominantly Latino ones. The coefficient for African American students within schools (the intercept for Column 3, Table 4) is substantially lower than the analogous coefficient from the first HLM (nine points rather than 16 points). Indeed, the coefficient drops to the level of statistical insignificance. The Latino coefficient (intercept for Column 4, Table 4) also changes substantially (from –9 to 27), and is statistically insignificant.[6] Thus, by including the 20 instructional practices, the second HLM can explain away the entire within-school racial gap.

The practices that reduce the gap seem to be somewhat different for African American and Latino students. Column 2 reveals some practices that are beneficial to all students, irrespective of race. Time on task is important; fourth graders who spend more time on math performed better on the assessment. Conducting routine exercises also proved helpful across the board, with a six point advantage to students whose teachers emphasized this. And of the topics, geometry proved the most beneficial. A few practices proved detrimental across the board. Frequent testing actually reduced scores on NAEP; working on projects and emphasizing facts (over reasoning and communication) also reduced scores. The practices particularly beneficial to African Americans and Latinos differed somewhat from those beneficial across the board and between the two ethnic groups. Beneficial practices are those with negative coefficients, as they reduce the racial gap, and detrimental practices are those with positive coefficients because they increase the gap. Thus, for black students the most beneficial practice is the emphasis on topics of measurement and estimation. On the other hand, testing has a disproportionately negative impact on black students, six points above and beyond the three points for all students. For Latino students, the most beneficial practice is the emphasis on data analysis. There are no practices analyzed here that proved specifically detrimental to Latino students.

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

[5] The coefficient is insignificant because even though the effect size is larger, the degrees of freedom are sharply reduced by including so many school-level independent variables.

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. Third, the list of practices studied is by no means comprehensive. As a result, the particular practices found to be associated with the achievement gap are suggestive that there is a broader range of practices that might influence the gap. The most important finding is not that the particular practices found to close the gap do so, but that what teachers do in the classroom as a whole makes such a difference to the gap.

Thus, the first conclusion is that instructional practices can affect the within-school achievement gap but not the between-school achievement gap. At a given school, depending upon what practices to which they are exposed, minority students will either keep up with their white peers or fall behind. This can be construed as a positive message in that schools really do have power; by emphasizing certain forms of instruction, school administrators can indeed succeed at closing the racial achievement gap in their schools. The flip side of the coin, however, is that a poor set of choices can either perpetuate or even worsen the achievement gap in a school, the black and Latino students falling behind year after year because the school does not know how to reach them. Second, the specific practices that seem to make the most difference are in specific topic areas. African American students appear to be particularly weak in measurement and estimation, and Latino students in issues around working with data. The other side of this coin is that school administrators need to redouble their efforts to provide solid instruction in these areas. The bulk of class time for fourth graders goes to numbers and operations. Yet teachers emphasizing these most basic topics is of no benefit to any demographic group. Instead, teachers need to spend more time on the topics they now spend the least time on, including geometry as well, insofar as that topic seems to benefit fourth graders across the board. Also, while it is not a race-specific finding, it should be emphasized here that the amount of time teachers devote to math is very important. With all of the emphasis in the early grades on reading, mathematics can fall through the cracks to some extent. Those teachers who are devoting below average amounts of time to math (less than 2.5 hours) would be well-advised to increase the amount of time. This said, the caveat mentioned above applies, that these practices may be indicative of a larger set of practices that influence the achievement gap.

From a policy perspective, these findings suggest a potential strength and a potential weakness of NCLB's mandate to close the racial achievement gap. To close the within-school portion of the gap seems eminently feasible with the types of interventions and accountability structure put into place. NCLB supports significant professional development in reading, mathematics, and science, and the placing of a qualified teacher in every classroom. If these goals are realized, principals should have at their disposal a corps of teachers that can not only raise achievement for all students at their schools, but can provide special attention to minority students. With sufficient attention, this study shows, any gaps within a given school can be completely eliminated. The accountability structure, by being primarily school based, reinforces this goal. It is schools, not districts, that must demonstrate AYP for all demographic groups. Principals thus have a strong incentive to institute instructional practices that will close the gap. Recent research has revealed that most school administrators do not believe that it is possible to achieve the NCLB mandate (Farkas et al., 2003). But rather than looking at the goal of schoolwide proficiency by 2013 as a non-starter, principals should see that it is in their hands to reduce racial inequality in their schools.

The way in which principals are powerless is in reducing racial inequality between their schools. Even if all students at all schools are making AYP targets, and within each school all students are performing at that higher level irrespective of race, schools with high minority populations are simply not going to meet the goal of proficiency by 2013. The policy instruments of increased professional development and accountability do not speak to the racial divide between schools; other policies are required. Perhaps equalizing resources between high and low minority schools would do the trick. Or the Federal Government or states could require greater racial balance between schools. Since residential patterns have become increasingly segregated, this would involve some inter-district form of busing. Or, if the core of the problem is the social isolation of inner-city schools and their surrounding communities, the Federal Government could provide community development assistance to neighborhoods that contain failing schools. Many options are possible, but research is required to know which would be the most effective in reducing the between-school gap, and then there would have to be political support for what could amount to an extremely expensive policy. But, as this study indicates, if policymakers want to attain the dream of NCLB, they are going to have to move beyond the current range of policy instruments contained in the current version of the ESEA.

References

Bell, L.I. (2003). Strategies that close the gap. *Educational Leadership, 60*(4), 32-34.

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.

Education Commission of the States. (2003). Retrieved November 22, 2004 from http://nclb.ecs.org/nclb/.

Farkas, S., Johnson, J., & Duffett, A. (2003). *Rolling up their sleeves: Superintendents and principals talk about what's needed to fix public schools.* New York: Public Agenda

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, D.C.: 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, D.C.: 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.
Haycock, K. (2003). Toward a fair distribution of teacher talent. *Educational Leadership, 60,* (4), 11-15.

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, D.C.: The Brookings Institution.

Jencks, C & Phillips, M. (1998). Introduction. C. Jencks & M. Phillips (eds.) *The Black-White Test Score Gap.* Pp. 1-51. Washington, D.C.: 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. Considerations and techniques for the analysis of NAEP data. *Journal of Education Statistics, 14,* 4, 303-334.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress Toward Equity? *Educational Researcher,31*(1), 3-12.

Lee, V.E. & Burkham, D.T. (2002). *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School.* Washington, D.C.: 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. (2002). *Embracing race: Whey we need race-conscious education policy.* New York: Teachers College Press.

National Center for Education Statistics. (2000). *NAEP 1999: Trends in academic progress.* Washington, D.C.: U.S. Department of Education.

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

Olson, L. (2003). In ESEA wake, school data flowing forth. *Education Week* (December 10):1,16-18.

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, D.C.: 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, D.C.: 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). Retrieved November 22, 2004 from http://epaa.asu.edu/epaa/v11n19/.

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives, 10*(12). Retrieved November 22, 2004 from http://epaa.asu.edu/epaa/v10n12/.

About the Author

Harold Wenglinsky
Hunter College School of Education, Room W1006
695 Park Ave
New York, NY 10021

E-mail: hwenglin@hunter.cuny.edu
Phone: (212) 772-4664

Harold Wenglinsky is currently an Associate Professor at the Hunter College School of Education. After receiving his Ph.D. in sociology at New York University, he was awarded a postdoctoral fellowship by Educational Testing Service to study the National Assessment of Educational Progress. He subsequently became a research scientist and then Director of the Policy Information Center, a research think tank housed at ETS. He worked at ETS for seven years before joining the Baruch College faculty in 2002. He is the author of numerous nationally recognized publications and has been the principal investigator of numerous projects, including two funded by the National Science Foundation. His primary expertise is in the analysis of large-scale databases to address issues of educational policy and practice, such as the roles of teachers and families in schools.

Education Policy Analysis Archives http://**epaa.asu.edu**

Editor: Gene V Glass, Arizona State University

Production Assistant: Chris Murrell, Arizona State University

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Gene V Glass, glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-2411. The Commentary Editor is Casey D. Cobb: casey.cobb@uconn.edu.