A Metaregression Analysis of the Effects of School Racial and Ethnic Composition on K–12 Reading, Language Arts, and English Outcomes

Roslyn Arlin Mickelson[1], Martha Cecilia Bottia[1], and Savannah Larimore[2]

Abstract

Racial and ethnic differences in educational outcomes significantly narrowed during the 1970s and 1980s when K–12 public schools were desegregated. However, when schools resegregated starting roughly in the late 1980s, racial gaps in outcomes widened again. Because of literacy's pivotal role in learning, the authors investigate if segregation contributes to racial gaps in K–12 reading performance. Drawing upon structural vulnerability and cumulative advantage/disadvantage theories to frame this study, the authors conduct multilevel metaregression analyses of 131 effect sizes from 30 primary studies to investigate if school composition effects contribute to racial gaps in K–12 reading outcomes and if any effects vary in magnitude or direction for students from different racial/ethnic backgrounds or grade levels. The metaregression analyses control for the primary studies' regression model characteristics and research designs. The results indicate a small, negative, statistically significant relationship between the percentage of a school's disadvantaged minority enrollment and the mean reading achievement of the students who attend it. The negative association is stronger when segregation is measured by percentage Black and is stronger for high school students. These two findings suggest that the disadvantages of segregated education cumulate as more structurally vulnerable students transition from elementary to secondary school. Additional results suggest that a school's racial composition effect is not the same as its socioeconomic status composition effect. The two organizational characteristics have distinct, albeit interrelated, influences on reading scores. Together the findings suggest that racially and ethnically segregated schooling both reflects and helps reproduce racial/ethnic inequality in literacy outcomes.

Keywords

multilevel metaregression, ethnic inequality, racial inequality, school segregation, K-12 reading achievement

School racial composition was reintroduced into the public conversation during the 2019 Democratic Party presidential candidate debates when Senator Kamala Harris critiqued former vice president Joe Biden's efforts decades ago to end mandatory busing for desegregation. But the issue was never really busing, which has historically been and continues to this day to be widely used to transport pupils from their homes to schools. The actual

[1]Department of Sociology, University of North Carolina at Charlotte, Charlotte, NC, USA
[2]Department of Sociology, Washington University in St. Louis, St. Louis, MO, USA

Corresponding Author:
Roslyn Arlin Mickelson, University of North Carolina at Charlotte, Department of Sociology, 9201 University City Boulevard, Charlotte, NC 28223, USA
Email: roslynmickelson@uncc.edu

issue then and now is the desegregation of racially and ethnically segregated public schools (Hannah-Jones 2019). The progress toward racial desegregation that commenced beginning in the late 1960s has faltered since the late 1980s, and many of the gains toward desegregation have reversed (Fiel and Zhang 2019; Logan, Minca, and Adar 2012; Reardon and Owens 2014; Stroub and Richards 2013). By many metrics, public schools in many areas are nearly as segregated as they were five decades ago (Frankenberg et al. 2019).

Notably, racial differences in achievement significantly narrowed during the decades when public schools were most desegregated. However, as school systems resegregated, racial gaps in virtually every outcome widened again (Berends and Peñalosa 2010; Bohrnstedt et al. 2015; Musu-Gillette et al. 2017). Is school segregation a driver of other racial and class stratification dynamics or a manifestation of them? More pointedly, does school segregation influence school outcomes?

Although answers to both questions are central for understanding education's role in generating and maintaining a racialized social order, we focus on the last question because to answer the first one, we need to know if segregation contributes to racially correlated achievement. We conduct a multilevel metaregression analysis of findings from studies examining school racial composition and K–12 reading, language arts, and English outcomes (henceforth, reading outcomes) conducted during the past 25 years. We chose to examine reading outcomes because of literacy's pivotal role in other educational processes and outcomes. Although reading outcomes are the product of numerous interacting individual, family, community, and classroom dynamics, we investigate one particular school structural characteristic—the racial and ethnic (henceforth, racial) compositions—of the schools that students attend.

It is important to examine whether school racial composition is part of the matrix of structural forces that undergird racially correlated differences in reading outcomes. If segregation contributes to the gaps, pursuing desegregation may be worth the political costs. However, if segregation is not a factor in racially correlated reading outcomes, we may lament the immorality of segregated education but decide that pursuing diverse public schools may not be a useful strategy for narrowing gaps in performance.

Our goal in this article is to shed light on this problematic by synthesizing prior social science literature that investigated the relationship of school racial composition to reading achievement. Although previous narrative syntheses include reading achievement along with other subject matter outcomes, to date, no comprehensive literature synthesis has focused exclusively upon reading achievement in relation to school racial composition or conducted a multilevel metaregression analysis to estimate the effect size of school segregation as we do in this article.

SEGREGATED, DESEGREGATED, AND RESEGREGATED PUBLIC EDUCATION

Segregated schools were foundational to Jim Crow society. The Supreme Court's decisions in *Brown v. Board of Education* (1954), holding that de jure racially segregated public schools were inherently unequal, and later in *Green v. County School Board of New Kent County* (1968), ordering that segregation be dismantled root and branch, began unraveling the fabric of legally sanctioned segregation in schools. It took a civil rights movement, additional lawsuits, and struggles on multiple fronts by determined and courageous people to eventually desegregate most public schools. From the mid-1970s through the late 1980s, vast numbers of public school systems were desegregated by redrawing attendance zones, strategically locating new schools, busing students to paired schools, and/or instituting magnet programs with wide cross-racial appeal. Initial desegregation efforts were triggered by lawsuits, federal governmental administrative orders, or court decisions. Later, some school districts voluntarily elected to desegregate by race or socioeconomic status (SES). The most recent Supreme Court decision on K–12 desegregation, *Parents Involved in Community Schools v. Seattle School District No. 1* (2007), successfully challenged the constitutionality of voluntary desegregation plans in Louisville, Kentucky, and Seattle, Washington that used individual student race as a criterion in assignments.

The legal landscape has shifted since the peak decades of desegregation. Currently, the federal government has essentially withdrawn from the desegregation struggle. Only a handful of cases remain on the docket of federal courts, and a small number of desegregation lawsuits are working their way through state courts (e.g., *Cruz-Guzman v. State of Minnesota* 2019). With some notable exceptions, federal, state, and local policy actors have turned much of their attention to school choice and other market-inspired reforms rather than equity-oriented strategies such as desegregation.

TRENDS IN STUDENT DEMOGRAPHY AND SCHOOL RESEGREGATION

The challenges presented by resegregating public education and racially correlated academic performance must be considered in conjunction with striking transformations in the demography of communities and their schools. Today, both are more ethnically and racially diverse and socioeconomically stratified than five decades ago. In 1968, when public schools were still largely segregated, 80 percent of U.S. public school students were White, 14 percent were Black, 5 percent were Latinx, and 1 percent was Asian or Native American. In 2017, the student population in public schools was roughly 45 percent White, 29 percent Latinx, 15 percent Black, 6 percent Asian and Pacific Islander, 4 percent biracial, and 1 percent American Indian/Alaskan Native (NCES 2019). Student populations have increasing numbers of immigrants, too. Approximately one quarter of children younger than 17 have at least one immigrant parent (Bottia 2019). Consequently, the proportions of the student population from more advantaged racial/ethnic backgrounds who tend to score well in reading are shrinking relative to the proportions of students from less advantaged backgrounds who are less likely to perform well.

The spatial geography of school segregation has changed as well. Increasingly, families of color live in inner ring suburbs, while some prosperous Whites are repopulating central cities. These demographic shifts are fueled by the growth of income inequality and the emergence of school choice. Given residential segregation (Rothstein 2017) and wide use of neighborhood-based assignment plans in public education, most pupils are likely to attend schools with others from similar racial and SES backgrounds.

READING, RACE, AND SCHOOL COMPOSITION

Racial/Ethnic Differences in Reading Performance

It is essential that students learn to read by third grade because by the fourth grade, they turn from learning to read to reading to learn other subjects. The problem is the stark disparities in reading performance among youth from different racial, ethnic, and socioeconomic backgrounds. Results from the National Assessment of Educational Progress (NAEP) indicate that literacy gaps appear early and continue through high school. White and Asian 4th graders score higher on average than Black, Latinx, and American Indian/Alaskan Native youth. Results from 2015 NAEP reading assessments of 12th graders indicate that 46 percent of White and 49 percent Asian youth read at or above proficiency compared with 28 percent of American Indian/Alaskan Native youth, 38 percent of biracial students, 25 percent of Latinx students, and 17 percent of Black youth (NCES 2015). Compounding these racial/ethnic gaps in proficiency are SES differences. Middle-class youth perform better than low-income pupils.

Nonschool Sources of Racial Differences in Reading. Social scientists and educational researchers investigating the mechanisms that underlie these differences demonstrate that individual student characteristics and family financial, cultural, and social capital resources all contribute to reading performance (Condron 2009; Lareau 2011; Roscigno and Ainsworth-Darnell 1999). Community resources that contribute to the acquisition of literacy include safety and crime levels, neighborhood SES, social networks, and cultural norms that embrace education (Jencks and Mayer 1990; Reardon and Bischoff 2011; Saporito and Sohoni 2007).

School Racial Composition and Reading Outcomes

Nonschool factors alone are insufficient to account for the racial differences in reading outcomes. School characteristics, including teacher and administrator quality, material resources, curricula, and instruction, are central to the literacy process. Recent studies by Bohrnstedt et al. (2015), Owens (2018), and Reardon (2016) have reaffirmed the centrality of school and district socioeconomic composition to learning outcomes. Since the late 1980s, a preponderance of research has identified school racial composition as an organizational characteristic that also influences academic performance (Mickelson, Nkomo, and Wimberly 2012).

The general findings from the corpus of relevant research about reading achievement are that net of individual characteristics, family background, teacher and principal quality, and various other school resources, a school's racial composition has a relationship to the reading achievement of all students from kindergarten to high school. The preponderance of these studies report a negative relationship between higher percentages of

disadvantaged minority, Black, or Latinx pupils and various measures of reading performance (Angrist and Lang 2004; Bali and Alvarez 2004; Bankston and Caldas 1996; Benson and Borman 2007; Borman and Dowling 2010; Brown-Jeffy 2006; Chatterji 2006; Condron 2009; Condron and Roscigno 2003; Condron et al. 2013; Crosnoe 2005; Gamoran and An 2005; Glenn 2006; Goddard, Salloum, and Berebitsky 2009; Hoxby 2000; Johnson and Nazaryan 2019; McCathern 2004; Mickelson 2001, 2015; Page, Murnane, and Willett 2008; Pong 1998; Reardon et al. 2019; Reid 2016; Roscigno 1998; Rumberger and Willms 1992; Southworth 2010; Stone, Brown, and Hinshaw 2010; Tevis 2007).

The negative effects of concentrating Black or Latinx students at a school appear to have a stronger impact on disadvantaged minority students themselves (Angrist and Lang 2004; Alexander, Entwisle, and Olsen 2014; Bankston and Caldas 1996; Benson and Borman 2007; Mickelson 2001, 2015). The evidence also indicates a positive relationship between a racially heterogeneous school and reading achievement for almost all students (Bali and Alvarez 2004; Brown-Jeffy 2006; Cook 1984; Liu and Carbonaro 2008; Southworth 2010). However, a smaller body of literature finds that racial composition is not significantly related to reading outcomes net of SES composition and other factors (Armor, Marks, and Malatinszky 2019; Armor and Duck 2007, Armor and Watkins 2006; Berk 2003; Chubb and Moe 1990; Rumberger and Palardy 2005; van Ewijk and Sleegers 2010). We believe that it is important to determine if there is a relationship between reading performance and school racial composition and, if one exists, to clarify its direction and strength.

CONCEPTUAL FRAMEWORK AND RESEARCH QUESTION

Our examination of the relationship between school composition and racially correlated differences in reading achievement is informed by two complementary theoretical frameworks. Structural vulnerability theory proposes that inequitable educational outcomes emerge as organizational features of schools interact with students' individual characteristics. Students' achievement is shaped by the interplay between their individual characteristics and the organizational structure of the schools in which they learn (Alexander et al. 2014; Hallinan 1991; Sørensen 1987). A school's racial composition is one such organizational feature, along with

ability grouping, tracking, and disciplinary processes. Students' own race, gender, and social class backgrounds can mediate or moderate how these school structures either enhance or constrain learning opportunities. For example, Reardon et al. (2019) showed that racial segregation is harmful precisely because it concentrates disadvantaged minority youth in high-poverty schools, which on average are less effective institutions compared with lower poverty schools. Students from lower SES or underserved minority backgrounds are more vulnerable to poor-quality schools because they are less likely to have family members with the financial, cultural, human, or social capital who can serve as a "safety net" to compensate for the less effective educations received in racially segregated schools. Thus, students whose race or class makes them structurally vulnerable are less likely to achieve in educational environments rendered inequitable by racial segregation.

Cumulative advantage is recognized as a mechanism for generating inequality across any temporal process in which a favorable (or unfavorable) relative position contributes to the further production of relative (dis)advantage. Research on cumulative advantage/disadvantage (CA/CD) as an inequality-generating process exists in sociological literatures about neighborhood effects, work and careers, health, and education (DiPrete and Eirich 2006). The CA/CD framework proposes that an individual initially exposed to advantages (or disadvantages) will accumulate further (dis)advantages from continued exposure over time, magnifying small differences and making it difficult for individuals or groups that are "behind" at one point in time to catch up. In the case of learning, initial small differences grow larger over time because progression from each step to the next depends on attainment of satisfactory performance in the previous step. Prior empirical research has used CA/CD to understand the association between teacher quality, special education placement, or tracking and students' outcomes over time (Gamoran and Mare 1989; Kerckhoff and Glennie, 1999; Lee and Mamerow 2019; Lucas 2001; Mickelson 2015; Sanders and Rivers 1996). The CA/CD framework is also directly relevant to reading outcomes. Reading skills have a reciprocal relationship with a set of cognitive skills, such that as reading skills increase so do these other cognitive skills, which in turn increase reading ability (Bast and Reitsma 1998; Stanovich 1986). Early reading abilities become resources for subsequent improvements in reading, as well as for learning in

other subjects. CA /CD suggests that younger students who attend lower quality schools are less likely to become proficient readers, and their weaker reading skills will cumulatively disadvantage them each year they attend a lower quality school. Racially segregated schools are, with few exceptions, less effective institutions than racially diverse ones. Researchers who connect the CA/CD framework to structural vulnerability theory argue that CA/CD results from not just an individual or a groups' position at the point of origin but from the interaction of complex forces (Dannefer 2003; DeLuca, Clampet-Lundquist, and Edin 2016). CA/CD does not question the importance of individual action; rather it highlights the power of structural realities within which human agency must operate. Together, structural vulnerability and cumulative advantage frame this study's investigation of school composition as a context for the generation of differences in reading achievement over the course of students' K–12 educational trajectories. Three research questions that guide our synthesis arise from the structural vulnerability and CA/CD frameworks:

1. Does the corpus of social science research since the late 1980s indicate that school racial composition is a significant predictor of reading achievement among K–12 students net of individual characteristics, family background, and other school factors, including school SES composition?
2. If it is, what is the direction and size of that effect?
3. If there is an effect, is it the same for students who come from different racial and ethnic backgrounds or grades in school?

METHODS AND DATA

The research design of this synthesis is a multilevel metaregression analysis. We focus on the past 25 years of scholarly research about the relationship between reading achievement and school racial composition because prior to the late 1980s, much of the research on compositional effects suffered from issues that undermined the reliability and validity of findings. Earlier studies often assessed desegregation effects before the desegregation "treatment" was fully implemented; they were generally small, district-level studies; researchers used comparatively unsophisticated statistical analyses; and the studies' samples frequently experienced large attrition rates over the course of the

desegregation treatment (St. John 1975). In contrast, more recent studies use cutting-edge statistical tools (such as multilevel modeling) and representative national, state, or district data sets. Importantly, the "treatment"—the specific policy designed to desegregate the district—had been implemented over a longer time frame than those evaluated in the earlier studies (Bradley and Bradley 1977; Cook 1984; Mickelson 2008). In the following section, we present an operationalization of crucial constructs used in our database searches and the protocol for assessment of suitability for candidate studies and subsequently in decisions to include studies in the metaregression analysis.

Definitions of Key Variables

School Racial Composition. The key independent variable in our study is school racial composition. Researchers vary in how they operationalized school racial composition across the studies we identified. Their nominal labels include *desegregated, diverse, integrated, racially isolated*, and *segregated*. The criteria for designating a school as desegregated or segregated can vary by district and within a district over time and typically depend upon the district's overall racial and ethnic mix at the time of measurement, court-ordered standards, or school board policy choices about pupil assignment. Our metaregression analyses incorporate a measure of school racial composition operationalized as percentage Black, percentage Latinx/Hispanic, percentage minority, percentage students of color, and so on.

Reading Outcomes. Our key outcome of interest is measured by standardized assessments in the form of various school-administered tests of reading. Most studies we include operationalize reading outcomes as standardized test scores. Standardized tests are problematic for many reasons, but they are widely used as a measure of achievement in the studies we synthesize. A minority of studies reported composite achievement scores that included reading along with mathematics, social studies, and science performance.

SES. SES is a critical control variable largely because it is so highly correlated with race at both school and student levels (Lucas and Beresford 2010). Distinguishing SES compositional effects from racial compositional effects is important for answering our motiving research questions. All studies included in this metaregression analysis

controlled for student-level SES, and 24 of 30 primary studies also controlled for school-level SES, typically by capturing the percentage of the student body qualified for free or reduced-price lunch, a common but imprecise indicator of SES. We model whether our primary study controlled for the SES of students and schools. Doing so allows us to identify if school racial composition shapes reading outcomes net of school-level SES composition.

Database Searches

We used a complete but parsimonious approach to our literature searches to address as many of the potential threats to their validity and reliability as possible (Raudenbush, Rowan, and Kang 1991). From 2006 through 2019, we conducted systematic searches of electronic databases in education, social, and behavioral science for relevant studies about the effects of school composition on these outcomes. The databases included JSTOR, Psych Abstracts, Sociology Abstracts, Google Scholar, ERIC, Educational Research Complete, Academic Search Premier, Project Muse, National Bureau of Economic Research, and Dissertation Abstracts.

With respect to reading outcomes, the keywords used in the searches (with an OR and an AND option) were selected because of their relevance to the topic studied in this metaregression analysis. The terms for the key independent variable included *racial composition, school racial composition, ethnic composition, school composition*, and various nominal labels associated with the construct, including *minority composition, desegregation, integration, segregation, racial isolation*, and *diversity*. The terms for the key dependent variables included phrases that signify academic achievement (*performance, outcomes, scores, test scores, grades, GPA*) in English, reading, or language arts.

Inclusion and Exclusion Criteria

In the first stage of our assessment process, a prospective study's abstract was retrieved and reviewed to determine if the study actually addressed the topic of interest. On the basis of the information provided in abstracts, we obtained full articles, chapters, books, dissertations, paper presentations, and reports for further evaluation for suitability for inclusion in the synthesis. In the second stage of our assessment, we subjected potential studies to the following preliminary inclusion criteria:

- The study examined the relationship of school racial composition to reading achievement.
- The dependent variable was a score that measured reading achievement either as a reading standardized ability estimate based on item response theory scores, a reading scale score, a composite score that included reading achievement such as overall grade point average (GPA), or a composite measure of statewide standardized tests in reading.
- The students in the study's sample were enrolled in an elementary or a secondary school.
- The study was written in English.
- The study's author(s) used appropriate statistical tools given the nature of the research design and the structure of the data. By "appropriate statistical tools," we refer to statistical techniques that allow researchers to conduct a more precise analyses in which the relationship between student achievement and school racial composition may be mediated or moderated by other school, district, individual, or family factors, and when appropriate, the study's models accounted for the nested nature of the achievement data.

Coding Procedure

Fifty-seven relevant studies met these preliminary inclusion criteria. We coded the studies that met all five initial inclusion criteria according to a formal coding protocol we developed for the larger project from which this metaregression analysis is drawn. The categories included for coding were (1) identifying information (author, title, journal, date of dissemination), (2) publication status, (3) research design, (4) description of the data set, (5) sampling frame, (6) sample characteristics, (7) independent and dependent variables, (8) keywords, (9) analysis method, and (10) key findings. We reviewed each code in all 57 studies to ensure the accuracy of the coding. Interrater agreement on codes was 98 percent. We collaboratively resolved uncertainties in coding that primarily revolved around the designation of research designs or sampling frames.

Selection of Primary Studies

We then subjected the 57 studies to four final inclusion standards required for calculating an effect

size for each regression coefficient that would be meta-analyzed:

- The key independent variable was measured as percentage racial/ethnic minority rather than percentage White or Asian students in the school.
- We required the key dependent variable to be reading grades, a composite measure that includes reading, or a reading test score, but not a gain score. Gain score studies compare the differences in students' performance from one period to another. A gain score that is correlated with a measure of racial composition will reflect the effects of racial composition on changes in gain scores over a specific time period, instead of achievement. Gain scores' range of values is smaller than the range possible with standardized achievement test scores. Both the mean and variance of the population of gain score regression effects are likely to be different than those for the population of single-point-in-time regression effects, making a synthesis with both types of effect sizes problematic.
- The study reported findings at the student level rather than at the school level.
- The study provided descriptive statistics for all regression coefficients reported as findings. Studies that otherwise met the inclusion criteria were unusable without complete descriptive statistics. Metaregression analysis requires descriptive statistics for all possible effects sizes calculated in each study so that all regression coefficients can be standardized across studies. Some otherwise qualified studies presented separate regressions for Blacks and Latinx in two or more grade levels, but the author provided only means, n's, and standard deviations for the overall sample, not the subsamples by race and grade level. We contacted researchers with requests for missing descriptive statistics (means and standard deviations for their dependent variable and key independent variables, and the sample size for all of the different relevant regressions in each study). We eliminated otherwise qualified studies whose authors were unable to provide us with the necessary missing information.

Final Sample of Primary Studies

The final sample of 30 primary studies (identified in the references with an asterisk) had 131 regression coefficients that served as the effect sizes we meta-analyzed. Twenty of the studies used reading grades, standardized tests, or standardized scores derived using statistical methods from item response theory; three studies used GPA (a composite measure); and seven studies used other types of composite measures that include reading achievement as dependent variables. Sixty-seven percent of primary studies used reading standardized test scores, rather than a composite score as a dependent variable. The majority of the 30 primary studies used sophisticated statistical techniques for data analysis (typically multilevel models or fixed-effects econometric models) and included controls for many student, family, and school characteristics. More than one-third of the 131 coefficient effects came from regressions that used percentage Black in the school as the main independent variable.

Analytic Procedures

We conducted a series of two-level hierarchical linear modeling metaregression analyses of the 30 primary studies and their 131 regression effects (Raudenbush, Bryk, and Congdon 2008). Metaregression analysis is a special case of multilevel modeling applied to research syntheses (Becker and Wu 2007). We began the construction of the data set by identifying or creating standardized regression coefficients within the 30 qualified studies. Next, we transformed all standardized coefficients using Fisher's z transformation to create a more normal distribution of effects for use in subsequent modeling and summarization (Mickelson, Bottia, and Lambert 2013). We used the following formula:

$$z = 0.5\left[\log(1+r) - \log(1-r)\right].$$

The z-transformed standardized reading achievement coefficients served as our dependent variable (r denotes the standardized coefficients).

We then examined the various independent variables used across the primary studies for potential use as control variables in our metaregression analysis. We treated the primary studies' regression model characteristics as level I predictors. The

Table 1. Regression Model Characteristics in 131 Effect Sizes Among Primary Studies.

Regression Model Characteristic	Percent of Models with Specific Characteristic
Controlled for family's income	63%
Controlled for lagged student achievement	30%
Controlled for parents' education	50%
Controlled for school's socioeconomic status composition	76%
Controlled for teacher characteristics	41%
Methodology controlled for nested structure of educational data	79%
Percentage Black students as the independent variable	46%
Percentage Latinx students as the independent variable	17%
Percentage minority students as the independent variable	34%
Sample included all racial categories	72%
Sample included only Asian students	1%
Sample included only Black students	5%
Sample included only Black and Latinx students	2%
Sample included only Black and White students	3%
Sample included only Black, White and Latinx students	4%
Sample included only Latinx students	6%
Sample included only White students	7%

characteristics of the primary studies' research designs serve as level II predictors in our metaregression analysis. Our choices of level I and level II predictors were constrained by the scope of the primary studies' research designs and coefficients in their regression models.

For the first full model we estimate, we selected level I and level II controls with theoretical or methodological importance for our research questions. Consistent with structural vulnerability theory and CA/CD theory, we included controls that accounted for individual student characteristics (such as race and SES) and school characteristics (such as teacher characteristics and school SES composition).

Table 1 presents descriptive statistics for characteristics of the regression models that produced the effect sizes in the 30 primary studies in our analytic sample. Among the 17 regression model characteristics appearing in the 30 studies, the following 6 regression model characteristics serve as level I predictors in our synthesis. We chose them because of their theoretical significance for our motivating questions:

- a control for family income is included,
- a control for school's SES composition is included,

- a control for teacher characteristics is included,
- the percentage of Black students is the independent variable,
- the sample included only Black students, and
- the sample included only Latinx students.

All level I predictors were entered as group mean centered to create intercept values that were equivalent to within-study mean effect size values.

Importantly, because all effect sizes are nested within primary studies, we included level II control variables that capture variability in research designs among the 30 primary studies we synthesize. Table 2 presents descriptive statistics for the 14 design characteristics found among these primary studies. We chose four research design characteristics to serve as level II predictors because, again, they had important theoretical or methodological significance for our motivating questions. All level II predictors were entered as grand mean centered. The following research design characteristics were used as level II predictors in our initial full metaregression model:

- whether the dependent variable in the study was longitudinal,
- whether the study used state-level data,

Table 2. Research Design Characteristics of 30 Primary Studies.

Research Design Characteristics	Percent of Designs with Specific Characteristic
Categorical key independent variable	17%
Continuous key independent variable	83%
Dependent variable was composite score that included reading/English (such as GPA)	33%
Dependent variable was cross-sectional	57%
Dependent variable was longitudinal	43%
Dependent variable was reading test score	67%
National level data	57%
Published study	77%
Purposive sample of schools	3 %
Sample included elementary school students	40%
Sample included high school students	57%
Sample included middle school students	27%
School district level data	13%
State level data	27%

- whether the sample included high school students, and
- whether the study used a reading test score as the dependent variable.

Our analytic steps began with estimating an unconditional model (model 1), followed by two full models, each with a slightly different set of level I and level II controls. The initial full model (model 2) uses the level I and level II control variables described above as theoretically relevant to our motivating questions. We estimated an additional full model (model 3) with a somewhat different set of level I and level II controls. For model 3, we selected level I and level II controls that appeared in 50 percent or more of the primary studies in our analytic sample, and therefore represent the research design of the typical study. Level I predictors in model 3 are

- a control measure for family income,
- a control for school SES composition,
- a methodology that controlled for the nested structure of data,
- a control for parent's education, and
- a sample that included students from all races.

The following study characteristics were used as level II predictors in model 3:

- the dependent variable is cross-sectional,
- national-level data are used,

- the sample includes high school students,
- reading test scores are the dependent variable,
- the independent variable is continuous, and
- the study was published.

We conducted the analysis reported in model 3 to serve as a reliability check on our initial choice of level I and II predictors for model 2. Importantly, all metaregression analyses were weighted to account for the fact that sample sizes differ across studies. Weighted analyses take into consideration the sample size of the primary studies by incorporating the inverse of the sampling variance of the effect sizes into the analyses.

RESULTS

The results of our two-level hierarchical linear metaregression analyses appear in Table 3. For the unconditional model (model 1), we nested effect sizes within studies and included no level I or level II predictors. The weighted and z-transformed standardized reading coefficients served as the dependent variable. This model was constructed to estimate both the overall average effect size and the between- and within-primary-study variance components. The between-study variance component accounted for 77.8 percent of the variance among the z-transformed coefficients. The within-study variance component accounted for 22.2 percent of the variance among the z-transformed coefficients.

The effect size estimate of –.080 represents the average *z*-transformed value once the nesting within the primary study was considered. The average *z*-transformed value when converted back into the standardized β weight scaling was also –.080. Results of the unconditional model indicate that attending a racially segregated school has a statistically significant negative relationship to reading achievement.

Models 2 and 3 in Table 3 present full models with somewhat different level I and level II control variables. Model 2 controls for variables that are theoretically meaningful given prior research on segregation effects. Model 3 presents a model with level I and level II controls most frequently used in the majority of the 30 studies included in this metaregression analysis. In both full models, all level I predictor variables were entered as group mean centered to create intercept values that are equivalent to within-study mean effect size values. After including control variables in both models 2 and 3, the results still indicate that attending a segregated school is negatively related to reading achievement. Results in both models show that the relationship is moderated by the characteristics of the primary study's research design as well as the characteristics of that study's regression model.

Model 2 indicates that the overall effect size estimate from the weighted full model with theoretically relevant level I and level II control variables was –.067. The results indicate a significant negative coefficient of –.076 for studies that used the variable percentage Black in a school (instead of percentage minority or percentage Latinx) as a measure of racial composition. Interpreted together, the results suggest that the negative effects associated with school racial segregation on reading achievement are stronger when school segregation is measured by the percentage of Black students. As expected, controlling for family income in a regression model moderates the relationship between reading achievement and school racial composition. The positive coefficient associated with studies that controlled for family SES (.229) suggests that controlling for family background in part reduces the negative association between percentage minority concentration in a school and students' reading achievement.

The level II controls in model 2 reveal that the effect size for studies that used longitudinal dependent variables were less negative (.135) than those that used a dependent variable that was cross-sectional in nature. Studies that include a reading score as a dependent variable, rather than a composite measure of achievement, yield larger negative coefficients (–.152), suggesting that the negative association between school minority composition and reading performance is even stronger when reading outcomes are measured by themselves and not as part of a composite measure of achievement. Studies using statewide databases, rather than national samples, produce larger negative associations between reading outcomes and school minority racial composition (–.078). And studies with samples that included high school students were more negative (–.119) than those with only elementary and middle school samples.

Model 3 indicates that the overall effect size estimate from the weighted full model with more commonly used level I and level II control variables was –.078, again reflecting a significant negative relationship between school racial segregation and reading achievement. In addition, model 3 shows that studies using a diverse sample of students yield more negative results (–.015) than those studies that include homogeneous samples of Black, Latinx, or white students. As is true in the other models, the inclusion of family income as a control variable in a design moderates the negative impact of racial segregation and reading achievement (.221). Studies with samples that include high school students (instead of just elementary students or middle school students) have a larger negative association between racial segregation and reading achievement (–.115). Studies with reading test scores as dependent variables and those that use a cross-sectional dependent variable produce more negative associations (–.130 and –.124, respectively) than studies that use a longitudinal dependent variable. Importantly, whether or not a study was published did not significantly increase or decrease the association between racial composition and reading achievement.

Last, our metaregression results in both models 2 and 3 show that controlling for school-level SES has no significant moderating relationship with the association between racial segregation and reading achievement outcomes. To further test this finding, we conducted a metaregression analysis with a sample of only studies that controlled for school-level SES (*n* = 24). Results of the analysis of the subsample are consistent with results obtained with the larger sample of 30 primary studies (available upon request).

Our level I findings from both models 2 and 3 offer empirical support for arguments that school racial and SES composition are separate organizational features of a school's opportunity structure

Table 3. Predictors of Effect Size Magnitude in Weighted Unconditional and Full Multilevel Metaregression Models.

	Model One Unconditional	Model Two Full Model Predictors Based on Theory	Model Three Full Model Predictors Commonly Used
Intercept	−0.080***	−0.067***	−0.078***
	(.018)	(.013)	(.016)
Level I Predictors-Effect size characteristics within studies			
Controls for family income		0.229***	0.221***
		(.044)	(.050)
Controls for school SES		−0.005	−0.005
		(.016)	(.016)
Controls for teacher characteristics		−0.014	
		(.027)	
Percent Black as independent variable		−0.076***	
		(.012)	
Sample includes only Black students		−0.014	
		(.010)	
Sample includes only Latinx students		.001	
		(.015)	
Methodology controlled for nested structure of educational data			−0.004
			(.011)
Controlled for parents' education			0.045
			(.060)
Sample includes students from all races			−0.015*
			(.009)
Level II Predictors-Study characteristics			
Dependent variable of longitudinal nature		0.135***	
		(.048)	
State-level sample		−0.078***	
		(.024)	−.0115**
Sample included high school students		−0.119**	(.051)
		(.048)	−0.130**
Dependent variable was reading test score		−0.152**	(.064)
		(.070)	−0.124***
Dependent variable was cross-sectional			(.047)
			0.030
Published study			(.031)
			0.038
National-level data			(.024)
			0.022
Continuous key independent variable			(.053)

*p < .05. **p < .01. ***p < .001.

that can interact with student characteristics in the reading achievement process. After we control for school SES, results in both model 2 and model 3 still show a negative relationship between school racial segregation and reading achievement. These findings reinforce the importance of research on school compositional effects that include measures for both SES and racial composition.

Table 4. Predicted Values of Hypothetical Model 2-Based Effect Size Estimates.

Estimates	z Transformation	B
Overall effect size estimate	−0.067	−0.067
Weighted overall effect size estimate	−0.067	−0.067
Level I Predictors: Effect size characteristics		
Controls for family income	−0.141	−0.142
Black students only	−0.053	−0.053
Latinx students only	−0.068	−0.068
Level II Predictors: Study characteristics		
Longitudinal dependent variable	−0.144	−0.145

To facilitate the interpretation of our multilevel metaregression models, we present an additional table with effect size estimates or the predicted values of z-transformed level I and level II predictors (Table 4). The effect size estimates in Table 4 are based on estimates in model 2 (the theoretically based model presented in Table 3) when selected variables are at their mean values. To do this, we programmed the full nested regression model into a spreadsheet and then systematically varied the values of the predictors to arrive at model-based predictions for a series of "what if" scenarios when predictors take different values. Table 4 presents both the z-transformed and the β weight scaling values of predicted values of the effect size magnitudes in different scenarios. The results show that the two types of estimates are very similar given that Fisher's z transformation has very little impact on small values. The first row shows the overall estimate (unweighted), the second row gives the weighted overall estimate, and level I and II variable estimates appear in subsequent rows. The overall average effect size was derived by weighting the effect sizes according to the method for random-effects models (Rosenberg 2005). This value was –.067, illustrating that the sample size of the primary study, in this case, does not influence the magnitude of the effect size obtained in this sample of primary studies.

The results of the multilevel metaregression analysis of model 2 were then used to estimate the effect sizes that included each of the model predictors. In the "what if" scenario in which all effect sizes of regressions included a control for family income, the model yields effect sizes that are larger in absolute value (–.141) than the overall average size effect. In the scenarios in which the coefficient comes from regressions with samples of Black students only or with samples of Latinx students only, the models yield predicted values of effect sizes slightly lower for Black samples (–.053) and almost the same as the overall average size effect for Latinx samples (–.068). In the "what if" scenario in which all studies had a longitudinal dependent variable, the predicted values yield effect sizes that are larger in absolute value (–.144) than the overall average effect size. Importantly, in all of these "what if" scenarios, the predicted values of the overall estimate are still negative, consistently yielding findings that indicate a negative relationship between school minority concentrations and reading achievement.

DISCUSSION AND CONCLUSION

These findings allow us to address our research questions. The first question asks if school racial composition influences reading outcomes. We find that it does. Results from our metaregression analyses indicate a statistically significant relationship between school minority concentration (Black, Latinx, and Native American students) and reading achievement. The results also indicate that a school's racial composition is not the same as its SES composition; the two organizational characteristics are distinct features of the school's structure of opportunity.

The second research question follows from the first one and concerns the direction and size of the relationship between school racial composition and reading outcomes. Our unconditional model (model 1) indicates a statistically significant small negative relationship between a school's disadvantaged minority enrollment and mean reading performance in the school. Once we enter theoretically important control variables into the analyses (model 2), the negative effect of racial segregation on reading scores weakens but is still statistically significant. Similarly, when we enter the most

commonly used control variables into the analyses (model 3), the negative effect remains and actually approaches the coefficient's size in model 1.

Although the magnitude of the minority enrollment effect is not large in absolute terms (–.067), it is roughly equivalent in magnitude (but not necessarily in direction) to the effect size of other reading curricular or instructional reforms as measured by standardized test scores (Lipsey et al. 2012). Moreover, the small effect size is far from trivial in substantive terms. To clarify how an overall effect of –.067 translates into students' standardized test scores, suppose a school system has an average school minority enrollment of 40 percent and a between-school standard deviation of 10 percentage points. The standardized reading achievement test score is scaled using the same approach commonly used by well-known standardized tests such as the SAT, with a mean of 500 and a standard deviation of 100. Under these conditions, and under the admittedly untested but plausible assumption that the effects of racial composition are linear, the results of this study would lead us to expect a difference of approximately 6.7 points in average test scores between two schools that were 1 standard deviation apart in minority racial composition. We would predict a school with 50 percent minority composition to have an average score of 493.3 (6.7 subtracted from 500), a school with a 60 percent minority population to have an average score of 486.6 (13.4 subtracted from 500), and so on. Our point in offering the illustration is that an effect size of –.067 associated with racial segregation ought not be considered a trifling blip. If we acknowledge that standardized test scores are used for evaluating school and teacher quality, student track placement, grades, promotion and other important matters whose effects cumulate over the years, we can begin to appreciate how segregation effects manifested in test scores cast a long shadow on educational outcomes.

The third question we pose asks whether the relationship between school racial composition and reading achievement varies for students from different backgrounds and grade levels. Structural vulnerability and cumulative disadvantage theories predict that school racial segregation will interact with students' characteristics such that any effects will be more harmful for marginalized youth and will increase in magnitude over the trajectory of their educational careers. Our findings lend support to these expectations. Results in model 2 indicate that in studies using percentage Black as an independent variable, instead of percentage minority and/or percentage Latinx, findings are more strongly negative. This suggests that segregation has a stronger negative relationship with reading outcomes when minority composition is measured as the percentage of Black students and is consistent with structural vulnerability theory. However, we do not find consistent variation of effect size by a sample's racial composition. Coefficients for samples that include just Black students ($n = 5$) or just Latinx youth ($n = 6$) were not statistically significant, likely because we identified very few qualified studies that include these distinctive samples. Overall, our metaregression results in models 2 and 3 suggest two possible interpretations: a concentrated presence of Black students is more consequential for all students in a school as the percentage Black grows, or there is a stronger impact of segregation for Black and Latinx students who attend segregated schools.

Results in models 2 and 3 show that high school students experience stronger effects of school racial segregation than youth in early grades, a finding consistent with cumulative disadvantage theory. It is worth noting that we obtained similar results in our previous metaregression analysis of school racial composition and mathematics achievement (Mickelson et al. 2013). The comparability of the results from both metaregression analyses conducted with largely different primary studies suggests the reliability of these findings with respect to the direction, magnitude, and likely cumulative effects of school composition on reading outcomes.

Unpacking the precise mechanisms that underlie the negative association between racial segregation and reading outcomes is beyond the scope of this study. Nonetheless, the broader literature on the topic offers a variety of possible mechanisms that are well within the structural vulnerability theoretical framework. One of the most likely mechanisms involves differential resources. Segregated minority schools have fewer resources than more diverse or racially isolated White schools. Teacher quality arguably is the most important school resource for reading outcomes. A recent longitudinal study of North Carolina and Washington state revealed that disadvantaged students are less likely to have quality teachers under every definition of student disadvantage and teacher quality (Goldhaber, Quince, and Theobald 2018). The recent history of Charlotte-Mecklenburg Schools (CMS) ties resegregation directly to lower quality teachers. Jackson (2009) used the natural experiment afforded by the 2002 end of desegregation to

examine the relationship of changing school racial composition to teacher quality in CMS. The return of CMS to a neighborhood school-based assignment plan triggered the rapid resegregation of the district. Using districtwide data from before and after the end of desegregation, Jackson found that schools that resegregated after they experienced an influx of Black students also experienced decreases in various measures of teacher quality. Jackson concluded resegregation caused better qualified teachers to transfer to more racially and socioeconomically diverse suburban schools.

Just as the benefits of having a good teacher, or the harms of low-quality teachers, deepen over the course of a student's educational career (Chetty, Friedman, and Rockoff 2011), our results indicate the ill effects of the negative association between racial segregation and reading outcomes appear to compound as students move from elementary through high school. Our finding that the high school association with minority enrollment is stronger than it is with elementary school suggests precisely this dynamic: the disadvantages of learning to read in segregated schools appear to cumulate as students move through the grades. Compounding the structural disadvantages of lower teacher quality associated with segregated schools over time is the consistent finding that students who attend schools with high enrollments of disadvantaged minority peers are themselves likely to be structurally vulnerable to low-quality teaching because of their own racial and SES backgrounds.

Cautions, Limitations, and Caveats

Cautions. Although it may be reasonable for us to draw these inferences from the findings, we offer them with caution for two reasons. First, all statistics in our primary studies' tables are about groups, not individuals. We do not have any information about the variability of the effect sizes for individual students. Second, the interpretations are only indirectly supported by the data we have available, because we were not able to include studies that tracked the regression effects over time for the same cohort of students.

Limitations. Both the relatively modest number of effect sizes ($n = 131$) and the small number of primary studies ($n = 30$) restricted the possible number of level I effect size characteristics and level II study characteristics that could be modeled in this metaregression analysis. Although we coded many more characteristics for both levels, we were forced

to select a subset of characteristics to model at each level from the overall set of characteristics we had coded. It is likely that our efforts to elaborate upon the relationships between school composition and reading outcomes do not capture all the mediating or moderating factors at play. Future metaregression analyses that use more studies will permit more sophisticated models to be tested.

Another limitation is actually the source of the first one: the modest number of studies used in the metaregression analysis raises the possibility that our findings suffer from the influence of sampling error. Future metaregression analysis with a larger number of studies of reading outcomes are needed to make sampling error less likely. A greater number of primary studies and effect sizes will allow more stable estimates of the underlying relationship between school organizational characteristics and reading outcomes.

An additional limitation is that the primary studies we synthesized used only linear models of the associations between enrollments of minority students and reading performance. In the future, researchers may wish to examine curvilinear models to evaluate whether there are unique effects for hypersegregated Black and/or Latinx schools that extend beyond a simple linear trend. There is a fair amount of research indicating that hypersegregated schools offer far fewer opportunities to learn than more diverse or merely racially imbalanced schools.

Caveats. The first caveat concerns the limits of metaregression analysis. Our models do not permit an examination of optimal ranges of ethnic and racial diversity. Narrative syntheses, such as the one conducted by Mickelson and Bottia (2010) and others we cited earlier, showed that diverse schools within certain ranges are not only better learning environments than segregated minority schools but are comparable or in some cases superior to racially segregated White schools. This kind of more nuanced examination of school composition and reading outcomes is not possible in this metaregression analysis.

Most important, a metaregression analysis of reading test scores privileges a very narrow set of intellectual skills associated with the domain of language acquisition, reading comprehension, and language use. Moreover, the focus on literacy ignores crucial noncognitive outcomes such as intergroup relations, which the preponderance of research indicates desegregation can foster (Braddock and Gonzalez 2010; Pettigrew and Tropp 2006). As Albert Einstein purportedly said, "Not everything that counts can be counted, and not everything that can be counted counts."

Conclusions

Racially correlated differences in reading performance narrowed significantly during the nearly two decades when U.S. schools were most desegregated. Racial gaps reemerged as the schools resegregated, even though overall U.S. students' reading performance has improved slightly in recent years. Yet gaps have increased between some groups in some grades. The 29-point 12th grade Black-White reading gap in the 2015 NAEP is much larger than it was in 2002 (Musu-Gillette et al. 2017).

These trends alone do not necessarily offer presumptive evidence of a relationship between racial segregation and school performance. Our findings provide a link. Although they do not offer evidence of causality, they clarify school racial composition's likely contribution to producing these trends. Schools with high enrollments of disadvantaged minority youth are effective delivery systems for unequal opportunities for learning to read. Older students who attend racially segregated minority schools fall more behind their otherwise comparable peers who learn reading in more racially diverse schools. This claim is consistent not only with our results but with 2015 NAEP reading scores that show significantly larger race gaps among 12th graders than among 4th graders. Just as the advantages of having a good teacher cumulate over the course of a student's educational career, so do the disadvantages of segregated schooling. Moreover, because learning to read is essential to learning in other subject areas, the cumulative effects of segregation cast a long shadow over a broad spectrum of learning year after year.

School racial segregation appears to be both a driver and a manifestation of racial stratification in education. It contributes to effectively reproducing the educational disadvantages that racially differentiated reading performance reflects. Given literacy's centrality to all learning, this nation is unlikely to break the intergenerational perpetuation of racism and fear, to prepare youth for citizenship in a democratic and just multiracial/ethnic society, or to equip every child to fully participate in a globalizing high-tech economy if we do not again consider the racial composition of the public schools we provide for our children.

REFERENCES

Alexander, Karl, Doris Entwisle, and Linda Olsen. 2014. *The Long Shadow: Family Background, Disadvantaged Urban Youth and the Transition to Adulthood*. New York: Russell Sage.

Angrist, Joshua, and Kevin Lang. 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review* 94(5):1613–34.

*Armor, David, Gary Marks, and Aron Malatinszky, 2018. "The Impact of School SES on Student Achievement: Evidence from US Statewide Achievement Data." *Educational Evaluation and Policy Analysis* 40(4):613–30.

*Armor, David, and Stephanie Duck. 2007. "Unraveling Black Peer Effects on Black Test Scores." Public Policy Program, George Mason University, Fairfax, VA. Unpublished manuscript.

*Armor, David, and S. Watkins. 2006. "School Composition and Hispanic Achievement." Public Policy Program, George Mason University, Fairfax, VA. Unpublished manuscript.

*Bali, Valentina, and Michael Alvarez. 2004. "The Race Gap in Student Achievement Scores: Longitudinal Evidence from a Racially Diverse School District." *Policy Studies Journal* 32(3):393–415.

*Bankston, Carl, and Stephen Caldas. 1996. "Majority Black Schools and Social Injustice: The Influence of De Facto Segregation on Academic Achievement." *Social Forces* 75(2):535–55.

*Bankston, Carl, and Stephen Caldas. 2000. "White Enrollment in Nonpublic Schools, Public School Racial Composition and Student Performance." *Sociological Quarterly* 41(4):539–50.

Bast, Janwillem, and Pieter Reitsma. 1998. "Analyzing the Development of Individual Differences in Terms of Matthew Effects in Reading: Results from a Dutch Longitudinal Study." *Developmental Psychology* 34(6):1373–99.

Becker, Betsy, and Meng-Jia Wu. 2007. "The Synthesis of Regression Slopes in Meta-Analysis." *Statistical Science* 22(3):414–29.

*Benner, Aprile, and Robert Crosnoe. 2011. "The Racial/Ethnic Composition of Elementary Schools and Young Children's Academic and Socioemotional Functioning." *American Educational Research Journal* 48(3):621–46.

*Benson, James, and Geoffrey Borman. 2007. "Family and Contextual Socioeconomic Effects across Seasons: When Do They Matter for the Achievement Growth of Young Children?" WCER Working Paper No. 2007-5. Madison: Wisconsin Center for Education Research.

Berends, Mark, and Roberto Peñalosa. 2010. "Increasing Racial Isolation and Test Score Gaps in Mathematics: A 30-Year Perspective." *Teachers College Record* 112(4):978–1007.

Berk, Richard. 2003. "Closing the Black-White Achievement Gap: Instructional Strategies in Title I Schools." PhD dissertation, Department of Education, Vanderbilt University, Nashville, TN.

Bohrnstedt, George, S. Kitmitto, B. Ogut, D. Sherman, and D. Chan. 2015. "School Composition and the Black-White Achievement Gap." NCES 2015-018. Washington, DC: National Center for Educational Statistics. https://nces.ed.gov/nationsreportcard/sub ject/studies/pdf/school_composition_and_the_bw_ achievement_gap_2015.pdf.

*Borman, Geoffrey, and Maritza Dowling. 2010. "Schools and Inequality: A Multilevel Analysis of Coleman's Equality of Educational Opportunity Data." *Teachers College Record* 112(5):1201–46.

Bottia, Martha C. 2019. "Educational and Residential Segregation of Immigrants in the United States." *Poverty and Race Research and Action Council Report* 28(1):1–8.

Braddock, Jomills, and Amaryllis Gonzalez. 2010. "Social Isolation and Social Cohesion: The Effects of K–12 Neighborhood and School Segregation on Intergroup Orientations." *Teachers College Record* 112(6):1631–53.

Bradley, Laurence, and Gifford Bradley. 1977. "The Academic Achievement of Black Students in Desegregated Schools: A Critical Review." *Review of Educational Research* 47(3):399–449.

Brown v. Board of Education, 347 U.S. 483 (1954).

Brown-Jeffy, Shelly. 2006. "The Race Gap in High School Reading Achievement: Why School Racial Composition Still Matters." *Race, Gender & Class* 13(3/4):268–94.

*Caldas, Stephen, and Carl Bankston. 1997. "Effect of School Population Socioeconomic Status on Individual Academic Achievement." *Journal of Educational Research* 90(5):269–77.

Chatterji, Madhabi. 2006. "Reading Achievement Gaps, Correlates, and Moderators of Early Reading Achievement: Evidence from the Early Childhood Longitudinal Study (ECLS) Kindergarten to First Grade Sample." *Journal of Educational Psychology* 98(3):489–507.

Chetty, Raj, John Friedman, and Jonah Rockoff. 2011. "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood." Working Paper 17699. http://www.nber.org/papers/ w17699.

Chubb, John, and Terry Moe. 1990. *Politics, Markets and America's Schools*. Washington, DC: Brookings Institution Press.

*Coleman, Alexandra. 2016. "A Quantitative Analysis of School-Level Factors and Their Impact on the Racial Achievement Gap." Unpublished PhD dissertation, Department of Sociology, University of Wisconsin, Milwaukee.

Cook, Thomas. 1984. "What Have Black Children Gained Academically from School Integration? Examination of the Meta-Analytic Evidence." Pp. 9–45 in *School Desegregation and Black Achievement*, edited by T. Cook. Washington, DC: National Institute of Education.

Condron, Dennis. 2009. "Social Class, School and Non-school Environments, and Black/White Inequalities in Children's Learning." *American Sociological Review* 74(5):683–708.

Condron, Dennis J., and Vincent J. Roscigno. 2003. "Disparities Within: Unequal Spending and Achievement in an Urban School District." *Sociology of Education* 76(1):18–36.

Condron, Dennis, Daniel Tope, Christina R. Steidl, and Kendralin J. Freeman. 2013. "Racial Segregation and the Black/White Achievement Gap, 1992 to 2009." *Sociological Quarterly* 54(1):130–57.

*Cooley, Jane. 2007. "Desegregation and the Achievement Gap: Do Diverse Peers Help?" Department of Economics, University of Wisconsin– Madison. Unpublished manuscript.

Crosnoe, Robert. 2005. "The Diverse Experiences of Hispanic Students in the American Educational System." *Sociological Forum* 20(4):561–88.

Cruz-Guzman v. State of Minnesota, 892 N.W.2d 533 (Minn. Ct. App. [2017]).

Dannefer, Dale. 2003. "Cumulative Advantage/ Disadvantage and the Life Course." *Journal of Psychological Social Science* 58(6):327–37.

DeLuca, Stefani, Susan Clampet-Lundquist, and Kathryn Edin. 2016. *Coming of Age in the Other America*. New York: Russell Sage.

DiPrete, Thomas, and Gregory Eirich. 2006. "Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments." *Annual Review of Sociology* 32:271–97.

Fiel, Jeremy, and Yongjun Zhang. 2019. "With All Deliberate Speed: The Reversal of Court-Ordered School Desegregation, 1970–2013." *American Journal of Sociology* 124(6):1685–1719.

*Fram, Melissa, Julie Miller-Cribbs, and Lee Van Horn. 2007. "Poverty, Race, and the Contexts of Achievement: Examining Educational Experiences of Children in the U.S. South." *Social Work* 52(4):309–19.

Frankenberg, Erica, Jongyeon Ee, Jennifer Ayscue, and Gary Orfield. 2019. "Harming Our Common Future: America's Segregated Schools 65 Years after *Brown*." The Civil Rights Project. https://www

.civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/harming-our-common-future-americas-segregated-schools-65-years-after-brown/Brown-65-050919v4-final.pdf.

*Gamoran, Adam, and Brian An. 2005. "Effects of School Segregation and School Resources in a Changing Policy Context." Paper presented at the meeting of the International Sociological Association Research Committee 28 (RC28) on Social Stratification and Mobility, Los Angeles, CA, August.

Gamoran, Adam, and Robert Mare. 1989. "Secondary School Tracking and Educational Equality: Compensation, Reinforcement, or Neutrality." *American Journal of Sociology* 94:1146–83.

Glenn, William. 2006. "Separate but Not Yet Equal: The Relation between School Finance Adequacy Litigation and Black Student Achievement." *Peabody Journal of Education* 81(3):63–93.

Goldhaber, Dan, Vanessa Quince, and Roddy Theobald. 2018. "Has It Always Been That Way? Tracing the Evolution of Teacher Quality Gaps in U.S. Public Schools." *American Educational Research Journal* 55(1):171–201.

Green v. County School Board of New Kent County, 391 U.S. 430 (1968).

Goddard, Roger, Serena Salloum, and Dan Berebitsky. 2009. "Trust as a Mediator of the Relationships between Poverty, Racial Composition, and Academic Achievement: Evidence from Michigan's Public Elementary Schools." *Educational Administration Quarterly* 45(2):292–31.

Hallinan, Maureen T. 1991. "School Differences in Tracking Structures and Track Assignments." *Journal of Research on Adolescence* 1(3):251–75.

Hannah-Jones, Nikole. 2019. "It Was Never about Busing." *The New York Times*, July 14. https://www.nytimes.com/2019/07/12/opinion/sunday/it-was-never-about-busing.html.

*Herman, Melissa R. 2009. "The Black-White-Other Achievement Gap: Testing Theories of Academic Performance among Multiracial and Monoracial Adolescents." *Sociology of Education* 82(1):20–46.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." Working Paper No. 7867. http://www.nber.org/papers/w7867.

Jackson, Kirabo. 2009. "Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation." *Journal of Labor Economics* 27(2):213–56.

Jencks, Christopher, and Susan Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood." Pp. 111–85 in *Inner-City Poverty in the United States*, edited by L. Lynn and M.G.H. McGeary. Washington, DC: National Academy Press.

Johnson, Rucker, and Alexander Nazaryan. 2019. *Children of the Dream: Why School Integration Works*. New York: Basic Books.

*Kainz, Kirsten, and Lynne Vernon-Feagans. 2007. "The Ecology of Early Reading Development for Children in Poverty." *Elementary School Journal* 107(5):407–27.

Kerckhoff, Alan, and Beth Glennie. 1999. "The Matthew Effect in American Education." *Research in Sociology of Education and Socialization* 12:35–66.

*Konstantopoulos, Spyros. 2005. "Trends of School Effects on Student Achievement: Evidence from NLS:72, HSB:82, and NELS:92." SSRN Scholarly Paper No. ID 822350. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=822350.

Lareau, Annette. 2011. *Unequal Childhoods: Class, Race, and Family Life*. 2nd ed. Berkeley: University of California Press.

*Lee, Hedwig. 2007. "The Effects of School Racial and Ethnic Composition on Academic Achievement during Adolescence." *Journal of Negro Education* 76(2):154–72.

Lee, Se Wong, and Geoff Mamerow. 2019. "Understanding the Role Cumulative Exposure to Highly Qualified Science Teachers Plays in Students' Educational Pathways." *Journal of Research in Science Teaching* 56(10):1362–83.

Lipsey, Mark, Kelly Puzio, Cathy Yun, Michael A. Hebert, Kasia Steinka-Fry, Mikel W. Cole, Megan Roberts, Karen S. Anthony, and Matthew D. Busick. 2012. "Translating the Statistical Representation of the Effects of Educational Interventions into More Readily Interpretable Forms." Washington, DC: National Center for Educational Statistics. https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf.

Liu, Guo, and William Carbonaro. 2008. "Friendship Networks and Racial/Ethnic Differences in Academic Outcomes." Paper presented at the meeting of the American Sociological Association, Boston, MA, August.

Logan, John, Elisabeta Minca, and Sinem Adar. 2012. "The Geography of Inequality: Why Separate Means Unequal in American Public Schools." *Sociology of Education* 85(3):287–301.

Lucas, Samuel R. 2001. "Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects." *American Sociological Review* 106(6):1642–90.

Lucas, Samuel, and Lauren Beresford. 2010. "Naming and Classifying: Theory, Evidence, and Equity in Education." *Review of Research in Education* 34:25–84.

McCathern, Daniel. 2004. "The Relationship between PreK–5 and K–5 Elementary School Size and Student Achievement." PhD dissertation, Department of Education, University of South Carolina, Columbia.

*McNalley, Julianne. 2005. "We Value Diversity, but . . . : Academic Achievement of White, Middle Class Elementary Students in Segregated and Integrated Schools." PhD dissertation, Department of Sociology, University of Iowa, Iowa City.

*Mickelson, Roslyn Arlin. 2001. "Subverting Swann: First- and Second-Generation Segregation in Charlotte, North Carolina." *American Educational Research Journal* 38(2):215–52.

Mickelson, Roslyn Arlin. 2008. "Twenty-First Century Social Science Research on School Diversity and Educational Outcomes." *Ohio State Law Journal* 69:1173–1228.

*Mickelson, Roslyn Arlin. 2015. "The Cumulative Disadvantages of First- and Second-Generation Segregation for Middle School Achievement." *American Educational Research Journal* 52(4): 657–92.

Mickelson, Roslyn Arlin, and Martha C. Bottia. 2010. "Integrated Education and Mathematics Outcomes: A Synthesis of Social Science Research." *North Carolina Law Review* 87:993–1089.

Mickelson, Roslyn Arlin, Martha C. Bottia, and Richard Lambert. 2013. "Effects of School Racial Composition on K–12 Mathematics Outcomes: A Metaregression Analysis." *Review of Educational Research* 83(1):121–58.

Mickelson, Roslyn Arlin, Mokubung Nkomo, and George L. Wimberly. 2012. "Integrated Schooling, Life-Course Outcomes, and Social Cohesion in Multiethnic Democratic Societies." *Review of Research in Education* 36:197–238.

*Muller, Chandra, Catherine Riegle-Crumb, Kathryn S. Schiller, Lindsey Wilkinson, and Kenneth A. Frank, K. 2010. "Race and Academic Achievement in Racially Diverse High Schools: Opportunity and Stratification." *Teachers College Record* 112(4):1038–63.

Musu-Gillette, Lauren, Cristobal de Brey, Joel McFarland, William Hussar, William Sonnenberg, and Sidney Wilkinson-Flicker. 2017. "Status and Trends in the Education of Racial and Ethnic Groups 2017." NCES 2017-051. Washington, DC: National Center for Education Statistics. https://nces.ed.gov/pubs2017/2017051.pdf.

NCES (National Center for Education Statistics). 2015. "2015 Mathematics and Reading: Grade 12 Assessments." http://nationsreportcard.gov/reading_math_g12/_2015/.

NCES (National Center for Education Statistics). 2019. "Estimates of Resident Population, by Race/Ethnicity and Age Group: Selected Years, 1980 through 2010." http://nces.ed.gov/programs/digest/d10/tables/dt10_021.asp.

Owens, Ann. 2018. "Income Segregation between School Districts and Inequality in Students' Achievement." *Sociology of Education* 91(1):1–27.

Page, Lindsay, Richard Murnane, and John Willett. 2008. "Trends in the Black-White Achievement Gap: Clarifying the Meaning of Within- and Between-School Achievement Gaps." Working Paper No. 14213, National Bureau of Economic Research, Cambridge, MA. http://www.nber.org/papers/w14213.

*Palardy, Gregory, Rumberger, Russ, and T. Butler. 2013. "The Effect of High School Segregation on Cognitive and Non-cognitive Skills in the United States." Paper presented at the International Conference on Segregation, Immigration, and Educational Inequality, Ghent University, Ghent, Belgium, September.

Pettigrew, Thomas, and Linda Tropp. 2006. "A Meta-analytic Test of Intergroup Contact Theory." *Journal of Social Psychology* 90(5):751–83.

*Pong, Suet-ling. 1998. "The School Compositional Effect of Single Parenthood on 10th-Grade Achievement." *Sociology of Education* 71(1):23–42.

Parents Involved in Community Schools v. Seattle School District No. 1, 127 S. Ct. 2738 (2007).

Raudenbush, Stephen, A. S. Bryk, and R. T. Congdon. 2008. "HLM Version 6.06." Lincolnwood, IL: Scientific Software International.

Raudenbush, Stephen, Brian Rowan, and Sang Jin Kang. 1991. "A Multilevel, Multivariate Model for Studying School Climate with Estimation via the EM Algorithm and Application to U.S. High-School Data." *Journal of Educational and Behavioral Statistics* 16(4):295–330.

Reardon, Sean. 2016. "School Segregation and Racial Academic Achievement Gaps." *Russell Sage Foundation Journal of the Social Sciences* 2(5):34–57.

Reardon, Sean, and Kendra Bischoff. 2011. "Income Inequality and Income Segregation." *American Journal of Sociology* 116(4):1092–1153.

Reardon, Sean, and Ann Owens. 2014. "60 Years after Brown: Trends and Consequences of School Segregation," *Annual Review of Sociology* 40:199–218.

Reardon, Sean F., Ericka Weathers, Erin Fahle, Heewon Jang, and Demetra Kalogrides. 2019. "Is Separate Still Unequal? New Evidence on School Desegregation and Racial Academic Achievement Gaps." CEPA Working Paper No. 1906, Center for Education Policy Analysis, Stanford, CA. http://cepa.stanford.edu/wp19-06.

Reid, Jeanne. 2016. "Race/Ethnic Diversity and Language Development in the Preschool Classroom." Pp. 39–55 in *School Integration Matters*, edited by E. Frankenberg, L. Garces, and M. Hopkins. New York: Teachers College Press.

*Roscigno, Vincent J. 1998. "Race and Reproduction of Educational Disadvantage." *Social Forces* 76(3):1033–61.

*Roscigno, Vincent J. 2000. "Family/School Inequality and African-American/Hispanic Achievement." *Social Problems* 47(2):266–90.

Roscigno, Vincent J., and James W. Ainsworth-Darnell. 1999. "Race, Cultural Capital, and Educational Resources: Persistent Inequalities and Achievement Returns." *American Journal of Sociology* 72(3):158–78.

*Roscigno, Vincent J., and Martha L. Crowley. 2001. "Rurality, Institutional Disadvantage, and

Achievement/Attainment." *Rural Sociology* 66(2): 268–292.

*Roscigno, Vincent J., Donald Tomaskovic-Devey, and Martha Crowley. 2006. "Education and the Inequalities of Place." *Social Forces* 84(4):2121–45.

Rosenberg, Michael S. 2005. "The File Drawer Problem Revisited: A General Weighted Method for Calculating Fail-Safe Numbers in Meta-analysis." *Evolution* 59(2):464–468.

Rothstein, Richard. 2017. *The Color of Law: The Forgotten History of How Our Government Segregated America*. New York: Norton.

Rumberger, Russell W., and Gregory J. Palardy. 2005. "Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School." *Teachers College Record* 107(9):1999–2045.

Rumberger, Russell W., and J. Douglas Willms. 1992. "The Impact of Racial and Ethnic Segregation on the Achievement Gap in California High Schools." *Educational Evaluation and Policy Analysis* 14(4):377–96.

*Ryabov, Igor, and Jennifer van Hook. 2007. "School Segregation and Academic Achievement among Hispanic Children." *Social Science Research* 36(2):767–88.

Sanders, William, and June Rivers. 1996. "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement." Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Saporito, Salvatore, and Deenesh Sohoni. 2007. "Mapping Educational Inequality: Concentrations of Poverty among Poor and Minority Students in Public Schools." *Social Forces* 85(3):1227–53.

Sørensen, Aage. 1987. "The Organizational Differentiation of Students in Schools as an Opportunity Structure." Pp. 103–29 in *The Social Organization of Schools*, edited by M. T. Hallinan. New York: Plenum.

Southworth, Stephanie. 2010. "Examining the Effects of School Composition on North Carolina Student Achievement over Time." *Educational Policy Analysis Archives* 18:1–42.

Stone, Susan, Timothy Brown, and Stephen Hinshaw. 2010. "ADHD-Related School Compositional Effects: An Exploration." *Teachers College Record* 112(5):1275–99.

Stroub, Kori, and Meredith Richards. 2013. "From Resegregation to Reintegration: Trends in the Racial/ Ethnic Segregation of Metropolitan Public Schools, 1993–2009." *American Educational Research Journal* 50(3):497–531.

Stanovich, Keith E. 1986. "Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy." *Reading Research Quarterly* 21(Fall):360–407.

St. John, Nancy. 1975. *School Desegregation Outcomes for Children*. New York: John Wiley.

*Tevis, Temika. 2007. "African-American Students' College Transition Trajectory: An Examination of the Effects of High School Composition and Expectations on Degree Attainment." Unpublished PhD dissertation, Department of Sociology, The Pennsylvania State University, State College.

Van Ewijk, Reyn, and Peter Sleegers. 2010. "Peer Ethnicity and Achievement: A Meta-analysis into the Compositional Effect." *School Effectiveness and School Improvement* 21(3):237–65.

AUTHOR BIOGRAPHIES

Roslyn Arlin Mickelson is Chancellor's Professor and a professor of sociology, women and gender studies, and public policy at the University of North Carolina at Charlotte. Her interests include how school organizational structures, resources, and processes interact with students' race, gender, social class, and immigrant status to shape K–16 educational outcomes.

Martha Cecilia Bottia is a research associate professor in the Department of Sociology at the University of North Carolina at Charlotte. She received her PhD in public policy and her MA in sociology from the University of North Carolina at Charlotte and her BS and MS in economics from the University of the Andes in Bogota, Colombia. Her interests include immigrant, minority, and gender issues in educational opportunities; school racial and ethnic composition effects, pathways to science, technology, engineering, and mathematics; and drug policies.

Savannah Larimore is a postdoctoral research associate in the Department of Sociology at Washington University in St. Louis. She received her PhD in sociology from the University of Washington in 2019. Her research interests are broadly related to the social determinants of racial and ethnic health disparities in the United States. Currently, many of her research projects focus on maternal and child health, specifically the risk for low birth weight.