



Evidence-supported interventions associated with Black students' educational outcomes

Findings from a systematic review of research

Michelle R. Same

Nicole I. Guarino

Max Pardo

Deaweh Benson

Kyle Fagan

Jim Lindsay

American Institutes for Research

Two expert reviewers who are unaffiliated with the Regional Educational Laboratory Midwest and American Institutes for Research offered comments to a previous draft of this report. The authors revised the report based on their comments. Any remaining errors or omissions are the responsibility of the authors.

February 2018

This Regional Educational Laboratory Midwest report was developed under Contract ED-IES-17-C-0007 from the U.S. Department of Education, Institute of Education Sciences. The report has undergone peer review. The content does not necessarily reflect the views or policies of the Institute of Education Sciences or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This report is available on the Regional Educational Laboratory Midwest website at <https://ies.ed.gov/ncee/edlabs/regions/midwest>.

REL Midwest at
American Institutes for Research
10 S. Riverside Plaza, Suite 600
Chicago, IL 60606
312.288.7600
RELMidwest@air.org



Summary

Research shows racial differences in educational access and academic achievement (Annie E. Casey Foundation, 2017; Office of Civil Rights, 2016). For example, Black students have less access to high-level math and science courses, are more likely to receive one or more out-of-school suspensions, and are more likely to attend schools with higher concentrations of inexperienced teachers than their White counterparts (Office for Civil Rights, 2016). Likewise, data from the 2015 administration of the National Assessment of Educational Progress indicated statistically significant score gaps between Black students and their White peers in math and reading for grades 4 and 8 in nearly all states with available data. The national high school graduation rates among Black students also lags behind those of their White peers (National Center for Education Statistics, 2017). These data on differences between Black and White students continue to raise concerns among researchers and education stakeholders (Barton & Coley, 2010; Coleman et al., 1966; Wixom, 2015). These educators and policymakers are seeking potential levers that might improve educational opportunities and achievement among Black students (for example, Horowitz & Samuels, 2017).

The Midwest Achievement Gap Research Alliance (MAGRA) is one group that seeks to identify such levers. MAGRA is a research-practice partnership made up of researchers, policymakers, and practitioners who share an interest in improving educational opportunities and outcomes for Black students. To help them begin this work, MAGRA tasked the Regional Educational Laboratory Midwest (REL Midwest) with addressing this foundational question:

What interventions have been shown to be associated with improved academic achievement of Black students according to evidence tiers I (strong evidence), II (moderate evidence), and III (promising evidence) from the Every Student Succeeds Act (ESSA)?

To address the question, REL Midwest's literature review team scoured an extensive set of research databases to find studies that explicitly mention associations between interventions and improved academic achievement of Black students, regardless of associations with outcomes for other student subgroups. For this research review, an intervention was defined as a policy, a practice, a program, or other malleable factor. Outcome measures of interest included students' achievement on standardized tests in English language arts (ELA) and math, high school graduation rates, and high school dropout rates.

The review team screened 3,917 report abstracts to identify quantitative studies published since 2002 that included a focus on Black students and presented an indicator of statistical association between an intervention and one or more of the outcomes of interest. For those abstracts that described a study meeting these criteria, the project team attempted to obtain the full-text study report. Following removal of studies that did not meet screening criteria, 53 studies remained.

The literature review team then examined the 53 studies more closely to determine if they met criteria for evidence in tiers I, II, or III per the U.S. Department of Education’s nonregulatory guidance on the ESSA (U.S. Department of Education, 2016). These tiers are labeled *strong evidence*, *moderate evidence*, and *promising evidence*, respectively (table 1). The detailed examination of the 53 studies found that no studies provided strong or moderate evidence, and 22 studies offer promising evidence. The 22 studies showed statistically significant associations between 20 interventions and the academic achievement of Black students (two interventions were each supported by evidence from two studies).

Table 1. Tiers of evidence from the Every Student Succeeds Act

Tier	Finding	Design	Sample	Other
Tier I (strong evidence)	Statistically significant association between intervention and favorable outcome	Meets WWC standards without reservations (can detect causal associations)	2+ sites AND 350+ students or 50+ schools, classrooms, teachers	No other studies meeting WWC standards that show statistically significant association with unfavorable outcome
Tier II (moderate evidence)	Statistically significant association between intervention and favorable outcome	Meets WWC standards with reservations (can detect causal associations)	2+ sites OR 350+ students or 50+ schools, classrooms, teachers	No other studies meeting WWC standards that show statistically significant association with unfavorable outcome
Tier III (promising evidence)	Statistically significant association between intervention and favorable outcome	Correlational study with statistical control for selection factors (cannot detect causal associations)	No sample size criteria	No other studies meeting WWC standards that show statistically significant association with unfavorable outcome

Note: WWC is What Works Clearinghouse.

Source: Based on U.S. Department of Education (2016).

The list of interventions identified in this report may not be exhaustive. Only studies that explicitly mention Black students in the abstract, keywords, or descriptors were eligible for this review. Studies of other interventions that included separate analyses of Black and White subgroups but neglected to mention racial differences in the abstract, keywords, or descriptors were not identified in the searches.

The studies and interventions identified in this review have the following characteristics:

- One intervention was implemented at the state level, 3 at the district level, 5 at the school level, and 7 at the classroom level. The other 4 interventions either were implemented at the student level or were tangentially related to the education system.
- Interventions could show positive associations with multiple outcomes. Eleven of the 20 interventions (55 percent) were positively associated with ELA achievement, 13 (65 percent) were positively associated with math achievement, 2 (10 percent) were negatively associated with high school dropout rates, and 1 (5 percent) was associated with high school graduation rates.

- Five of the 20 interventions had positive associations for students in elementary school (grades K–5), 3 had positive associations for students in middle school (grades 6–8), 5 showed positive associations for students in high school (grades 9–12), and 7 showed positive associations for students at multiple school levels.

The report describes the interventions and the studies that provide the promising evidence.

Contents

- Summary i
- Why this research review?.....1
 - Racial disparities in educational opportunities and outcomes1
 - Increasing educational opportunities and outcomes among Black students is a priority in Wisconsin.....2
 - This research review examined the evidence on interventions that aim to improve Black students’ educational outcomes.....2
- What the review examined.....3
- What the review found7
 - Policy at the state level.....8
 - Local district policies9
 - Policies and practices in the school community.....9
 - Teacher practices in the classroom13
 - Supplemental interventions16
 - Interventions absent from this review19
- Implications19
 - The 22 interventions identified in this review need to be tested further.....19
 - Some interventions need better articulation20
 - The intervention roll-out strategy can help accumulate additional evidence20
- Limitations of the research review21
- Appendix A. Literature review methodology A-1
- Appendix B. Interventions supported by promising evidence.....B-1
- Appendix C. Interventions with no associations or negative associations with Black students’ outcomes..... C-1
- References Ref-1

Boxes

Box 1. Tiers of evidence according to the U.S. Department of Education’s nonregulatory guidance on the Every Student Succeeds Act (U.S. Department of Education, 2016)	4
Box 2. Review methodology	5

Tables

Table 1. Tiers of evidence from the Every Student Succeeds Act	ii
Table A1. Keywords and descriptors used in the literature search.....	A-4
Table B1. Interventions associated with Black student academic outcomes and characteristics of studies that provide promising evidence for those associations...	B-1
Table C1. Interventions showing no associations with Black students’ achievement outcomes or significant associations with unfavorable outcomes	C-1

Figures

Figure A1. The interventions with strong, moderate, or promising evidence were identified through a research review involving five steps	A-7
--	-----

Why this research review?

In the United States, Black children are more likely than their White counterparts to face certain barriers to educational success (Annie E. Casey Foundation, 2017a). For example, Black children are more likely than their White peers to grow up in poverty, have less access to fresh and healthy food, have greater exposure to environmental hazards, and have less access to human service programs such as Temporary Assistance for Needy Families (Bower, Thorpe, Rohde, & Gaskin, 2014; Collins, Munoz, & JaJa, 2016; Cushing et al.; McDaniel, Woods, Pratt, & Simms, 2017). Differences in academic achievement, or the achievement gap¹, between Black students and White students become apparent as children enter the K–12 school system (Jacobson Chernoff, Flanagan, McPhee, & Park, 2007; Musu-Gillete et al., 2017). The educational outcomes of Black students relative to their White peers continue to raise concerns among educators, policymakers, and the public at large (Barton & Coley, 2010; Coleman et al., 1966; Wixom, 2015).

Racial disparities in educational opportunities and outcomes

The challenges faced by Black students extend into the K–12 system. Nationally, Black students have less access to high-level math and science courses, are 3.8 times as likely to receive one or more out-of-school suspensions, and are more likely than their White peers to attend schools with higher concentrations of inexperienced teachers and teachers who have not met all state certification and licensure requirements (Office for Civil Rights, 2016).

Data from the 2015 administration of the National Assessment of Educational Progress (NAEP) indicated statistically significant score gaps were present between Black students and their White peers in math for grades 4 and 8 in all states with available data (that is, in 42 and 39 states, respectively, by grade). NAEP data from that same year also showed statistically significant gaps in reading for grades 4 and 8 in all but one state with available data (that is, in 40 of 41 and all 42 states for grades 4 and 8, respectively; National Center for Education Statistics, 2017).

Differences in achievement between Black students and White students also include high school graduation and dropout rates. At a national level, Black students graduate at lower rates than do their White peers. In the 2014/15 school year, 88 percent of White students and 75 percent of Black students graduated from public high schools within four years of starting grade 9 (National Center for Education Statistics, 2017). The gaps in educational attainment have prompted ongoing research and debate on potential levers that can improve educational opportunities and achievement among Black students (for example, Horowitz & Samuels, 2017).

¹ Some scholars argue that the achievement gap should be reframed as an opportunity gap (for example, Carter & Welner, 2013). This report uses the term *achievement gap* to be consistent with observable accountability measures.

Increasing educational opportunities and outcomes among Black students is a priority in Wisconsin

While achievement gaps between Black students and their White peers exist in every state for which data are available, the gaps among Wisconsin students have consistently been among the largest in the United States. Wisconsin exhibited the largest achievement gap between Black students and their White peers on the 2015 NAEP math and reading tests for both grades 4 and 8 (National Center for Education Statistics, 2017). Likewise, in 2015, Wisconsin also held the largest school graduation rate gap between Black students and White students in the nation (Center for Education Statistics, 2016).

Increasing educational outcomes among Black students has become one of the Wisconsin Department of Public Instruction's (DPI's) highest priorities (Wisconsin Department of Public Instruction, 2014). Accordingly, leaders from the DPI, the Regional Educational Laboratory Midwest (REL Midwest), and other stakeholder groups in Wisconsin have formed a research-practice partnership, called the Midwest Achievement Gap Research Alliance (MAGRA).

To aid their work in increasing educational outcomes among Black students, MAGRA expressed an urgent need for information on interventions that are associated with the academic achievement of Black students. The alliance requested that REL Midwest fulfill this need by conducting a search of the research literature to identify such interventions. The alliance plans to build off this work by examining whether districts and the state education agency are implementing such interventions. If not, then REL Midwest and its MAGRA partners can provide training, coaching, and technical support to districts and the state education agency on interventions that may increase the achievement of Black students.

This research review examined the evidence on interventions that aim to improve Black students' educational outcomes

For this research review, an intervention was defined as a policy, a practice, a program, or other malleable factor, regardless of whether it is enacted at the state, district, school, teacher, classroom, or student level. Operationally, school systems can affect the Black-White achievement gap by offering interventions to poorly performing students, schools, or districts. To the extent that Black students (or schools or districts serving mostly Black populations) make up a disproportionate percentage of these poorly performing populations, then assigning members of these poorly performing populations to evidence-supported interventions should move the achievement of Black students closer to that of their White counterparts. This research review aimed to identify interventions that are associated with Black students' academic outcomes, regardless of associations with outcomes for other student subgroups.

The specific outcomes of interest include student achievement based on assessments in English language arts (ELA, including reading, writing, or literacy) and math as well as high school graduation and dropout rates. To be eligible for this review, a study report had to include a statistic that indicates an association (that is, correlation)² between an intervention and at least one of these key outcomes. For student achievement in ELA and math and high school graduation rates, studies had to show that interventions were positively associated with the outcome for Black students. For high school dropout rates, studies had to show that interventions were negatively associated with the outcome for Black students.

What the review examined

This review addressed the following research question:

What interventions have been shown to be associated with improved academic achievement of Black students according to evidence tiers I (strong evidence), II (moderate evidence), and III (promising evidence) from the Every Student Succeeds Act (ESSA)?

To be classified as supportive evidence, studies had to meet the standards for these tiers from the ESSA, as articulated in the U.S. Department of Education’s nonregulatory guidance (U.S. Department of Education, 2016; see criteria for tiers in box 1).

To address the research question, REL Midwest searched research databases to find all studies that focused on Black students’ academic achievement or the Black-White achievement gap and were published in 2002 or later.³ The literature review team screened study abstracts and full text reports to eliminate studies that did not meet the eligibility criteria (listed in box 2). Studies that survived the screening process were reviewed more thoroughly to determine whether they met the criteria for strong, moderate, or promising evidence (see box 1; box 2 provides a general overview of the research review’s methodology, with more details provided in appendix A).

² A correlation is a statistic that indicates the relationship between two variables. Positive correlations indicate that as values for one variable increase, so do values for the other variable. Negative correlations indicate that as values for one variable increase, values for the other variable decrease. Statistical comparisons between groups operate in the same way, except one variable is binary (that is, the intervention was present or not present). Most statistical procedures that examine the relationships between variables and impacts of interventions are based on the correlation coefficient, but the actual names of the statistics and their metrics vary. Therefore, the phrase *associated with* is used throughout the report to indicate the relationship between an intervention and an outcome measure.

³ The decision to limit the time frame to 16 years was based on MAGRA’s preference to identify interventions that have been studied during the years in which state and local education agencies have worked to reduce Black-White achievement gaps (as mandated by the No Child Left Behind Act of 2001).

Box 1. Tiers of evidence according to the U.S. Department of Education’s nonregulatory guidance on the Every Student Succeeds Act (U.S. Department of Education, 2016)

- **Tier I (strong evidence):** The study must be a well-designed and well-implemented randomized controlled trial (RCT) with low attrition (that is, meets What Works Clearinghouse [WWC] standards without reservations), include a large sample (at least 350 students or 50 or more groups of students) across multiple sites (more than one local education agency, locality, or state), show a statistically significant favorable finding, and have a sample that overlaps with the population and setting proposed to receive the intervention. There also must be no studies of the same intervention that meet tier I or II design standards but show a statistical association with an unfavorable outcome.
- **Tier II (moderate evidence):** The study must be a well-designed and well-implemented RCT with high attrition or a quasi-experiment, with an intervention and a comparison group equivalent at baseline (that is, meets WWC standards with reservations); have a large, multisite sample (as defined for tier I); show a statistically significant favorable finding; and have a sample that overlaps with the population or setting proposed to receive the intervention. There also must be no studies involving the same intervention that meet tier I or II design standards but show a statistical association with an unfavorable outcome.
- **Tier III (promising evidence):** The study must be a well-designed and well-implemented correlational study that statistically controls for selection bias and must show at least one statistically significant association between the intervention and a favorable student outcome. RCTs and quasi-experimental design studies that did not meet criteria for strong or moderate evidence still could be classified as tier III, if statistical models controlled for selection bias. There also must be no studies involving the same intervention that meet design standards for tiers I–III but show a statistical association with an unfavorable outcome.

Note: According to the ESSA, a study also can be classified as providing tier IV evidence if it includes an evidence-based rationale for why statistical associations should exist between interventions and outcomes. Studies providing tier IV evidence were not included in this review.

Box 2. Review methodology

The research review team employed a three-phase process to ensure that the review was thorough and systematic. This box provides an overview of the three phases (appendix A provides more details).

Phase 1: Search of literature databases

Research studies on interventions that might affect Black students' achievement were identified by searching two online databases: the Education Resources Information Center and Google Scholar. An additional eight research literature databases were searched through the EBSCO platform (for example, Academic Search Premier, Education Source, and PsycINFO). Search terms entered into the literature databases included three components:

[relevant achievement outcomes] AND [relevant grade level bands] AND [sample focus terms: Black-White achievement gap, Black students, African American, and others].

For keyword searches of databases, the terms relevant to each component were entered and connected with OR. The resulting search included all combinations of outcomes of interest, grade levels, and terms related to Black students (see table A1 in appendix A for terms). The database searches identified 7,177 potential studies. Members of the Midwest Achievement Gap Research Alliance (MAGRA) also nominated 16 studies. After the review team removed 3,273 duplicates and 3 studies published before 2002, 3,917 studies remained.

Phase 2: Screening out ineligible studies

Screening criteria were reviewed with MAGRA members. The criteria were as follows:

Report type: technical reports and working papers from research organizations, doctoral dissertations, and articles in research or practitioner journals.

Dates of publication: between 2002 and June, 2017 (16 years).

Study sample: students in grades K–12 who reside in the United States.

Study design: randomized controlled trial, quasi-experimental design, or correlational study that statistically controls for student selection factors.

Outcomes: student achievement on standardized tests in English language arts (which includes reading, writing, and literacy) or math, high school graduation rates, or dropout rates.

Intervention: a policy, a practice, a program, or some other malleable factor.^a

Unique associations for Black students: one or more statistically significant associations for a sample consisting of at least 80 percent Black students^b or a Black subgroup of students.

Five trained screeners examined the 3,917 abstracts to determine whether they were eligible. The abstract screen removed 2,976 studies. A research librarian attempted to obtain the remaining 941 studies. She was unable to locate 234 studies, and 146 other studies could not be obtained in time to be included in the review.^c The remaining 561 full-text reports were then examined against the same criteria. After removing the 508 studies that did not meet the screening criteria, 53 studies remained. To monitor the accuracy and consistency of screeners' judgments, two review team members screened a 10 percent sample of abstracts and full texts.^d The principal investigator resolved any disagreements.

Phase 3: Collecting data from eligible studies and classifying evidence using the tiers from the Every Student Succeeds Act (ESSA)

The studies that passed through both screening phases were then assigned to reviewers who recorded the intervention name, the delivery type (for example, online or teacher led), the level of implementation (for example, school or classroom), the duration of implementation and the frequency of intervention sessions, outcome measures, statistical results, and features of the study sample and context (for example, school demographics and location). Finally, two team members, one of whom was certified as a reviewer for the What Works Clearinghouse (WWC) group design standards (version 3.0), classified the evidence in the studies as *strong* (tier I), *moderate* (tier II), *promising* (tier III), or not in tiers I–III according to the ESSA tiers of evidence (see box 1).^e This classification of studies eliminated 31 studies because they did not meet the criteria for tiers I–III.

Notes:

- a. Malleable factors are defined as factors that are alterable, such as children’s behaviors; technologies; and education programs, policies, and practices (see Earle et al., 2013).
- b. This criterion was adopted based on a review by Bandy and Moore (2011).
- c. The project team was unable to obtain 146 full texts given the project’s time frame and limited resources. These included 116 doctoral dissertations, 21 books, and 9 journal articles.
- d. Consistency rates were 89 percent for abstracts and 77 percent for full texts. Most of the low consistency rate for full text screening (87 percent of mismatches) reflected one screener’s preference for the “Don’t Know” response option. When screeners made this response to any full text report, the final disposition for report was determined through discussion between the primary investigator and the screener(s).
- e. To be classified as providing strong or moderate evidence (tier I or II), studies must meet WWC group design standards without reservations (for strong evidence) and WWC group design standards with reservations (for moderate evidence).

The 22 studies that passed through the process described in box 2 represent 20 interventions having at least one statistically significant positive association with an academic achievement outcome,⁴ with no studies showing statistically significant negative associations.

The 20 interventions and supporting evidence are described in the following section. Readers are cautioned against assuming that such associations will be found in all settings. These associations may represent average associations across numerous settings, and the associations for specific settings may vary.

Moreover, the interventions described in the following section are supported by evidence that was generated by individual studies. Other studies of the same interventions may exist that found no statistically significant associations with Black student outcomes. In such cases, the evidence rating was not adjusted to account for studies showing no statistically significant associations with Black student outcomes.⁵

⁴ For associations with high school dropout rates, a negative association indicates a favorable outcome.

⁵ See the limitations section for further discussion.

What the review found

The research review found 53 studies that met the eligibility criteria, and of those, 22 met the standards for strong, moderate, or promising evidence described in the U.S. Department of Education’s nonregulatory guidance on the ESSA (U.S. Department of Education, 2016). The studies provide evidence supporting 20 interventions. The following are characteristics of the studies and interventions that passed the screening and review processes:

- All 22 studies provided promising (tier III) evidence. None of these studies had the combination of research design and sample size that was necessary to be classified as strong or moderate evidence.
- Of the 20 interventions identified in these studies, 1 was implemented at the state level, 1 were implemented at the district level, 7 at the school level, and 7 at the classroom level. The other 4 were only partially implemented by school staff, were implemented outside of normal school hours, or could have been implemented anywhere.
- Interventions could show positive associations with multiple outcomes. Twelve of the 20 interventions were positively associated with ELA achievement, 11 were positively associated with math achievement, 2 were negatively associated with high school dropout rates, and 1 was positively associated with high school graduation rates.
- Five of the 20 interventions had positive associations for students in elementary school (grades K–5), 4 had positive associations for students in middle school (grades 6–8), 5 showed positive associations for students in high school (grades 9–12), and 6 showed positive associations for students at multiple school levels.

The design, the implementation, or the reporting of the studies was not sufficient to infer that the interventions caused the outcomes. Although statistical procedures can control for other potential causal factors, correlational studies reveal only statistical associations; they cannot determine the direction of influence between an intervention and outcome.

To help structure the presentation of findings, the interventions for which promising evidence was found are presented according to the level of education system in which they were implemented. The levels are labeled as follows:

- Policy at the state levels
- Local district policies
- Policies and practices in the school community
- Teacher practices in the classroom
- Supplemental interventions

The last level—supplemental interventions—includes those that are only partially implemented by school staff, those that occur outside the normal school hours, or those that can be implemented anywhere.

Policy at the state level

District Assistance and Intervention Teams

To assist school districts that failed to meet their adequate yearly progress goals⁶ for four or more consecutive years, California implemented a two-tiered system of technical assistance. One tier involved districts with the lowest performing schools according to the most recent accountability tests, and these districts were determined to be most in need of moderate or intensive assistance. These districts received \$100,000 to \$150,000 per underperforming school to contract for two years with highly regulated, state-approved external experts known as district assistance and intervention teams (DAITs). DAITs worked closely with district leaders to assess district needs in nine essential program components and develop a capacity report that summarized the needs in the nine components. DAIT staff then worked with districts to write an improvement plan based on the recommendations from the capacity report. Then DAITs staff provided the districts with the human capital (that is, the knowledge of reforms and social ties to other organizations) to help leverage the reforms enumerated in the district's improvement plan. The other tier was for districts that failed to meet adequate yearly progress goals for four or more years but performed better according to the most recent accountability tests. These districts were determined to need less intensive technical assistance. Such districts received \$50,000 per underperforming school to contract with more traditional, less regulated technical assistance providers of their choosing. These non-DAIT technical assistance providers often addressed narrower, district-identified issues.

To explore the relationship between DAITs and student achievement, researchers used data from all California public school students in grades 2–11 from 2005/06 to 2010/11 for whom either ELA or math test data were available, resulting in 26.3 million student-year observations for the full sample (Strunk & McEachin, 2014). The researchers then employed a difference-in-difference approach to compare students in DAIT districts to students in non-DAIT districts before and after intervention. After controlling for school; district; time; student prior achievement; race; and English learner, free or reduced-price lunch, and disability status, analyses indicated that DAIT technical assistance was associated with positive increases in math achievement for Black students, relative to White students. This study was classified as offering

⁶ Adequate yearly progress was a term used in the No Child Left Behind Act of 2001 to refer to student achievement growth for schools and districts generally and among student subgroups within schools and districts. The definition of growth and the assessments used to measure growth were determined by states.

promising evidence because it was a correlational study with statistical controls for other student and district factors.

Local district policies

Hiring certified teachers

Two studies used data from the Early Childhood Longitudinal Study-Kindergarten (ECLS-K) to determine whether teachers' certification status was associated with Black students' reading achievement in grades K–3 (Easton-Brooks & Davis, 2009; Graves, 2011). These studies had different requirements for inclusion in the analytic sample,⁷ and the analyses included different sets of student, family, and teacher characteristics as control variables. Both studies found that Black students whose teachers held a “standard certificate” (defined as having a full certificate or an advanced certificate) showed significantly more reading growth between kindergarten and grade 3 than Black students whose teachers did not hold a standard certificate (defined as holding no certificate, holding a probationary certificate, or holding a certificate issued through an alternative program; Easton-Brooks & Davis, 2009; Graves, 2011). The evidence from these two studies was classified as promising because of each study's correlational design with statistical controls for student characteristics.

Policies and practices in the school community

The Elementary School Success Profile Model of Assessment and Prevention

Rather than being a specific intervention itself, the Elementary School Success Profile Model of Assessment and Prevention (ESSP MAP) is a tool that collects and organizes data to aid school administrators in better understanding the circumstances and needs of their students. Using data collected from online assessments administered to teachers, parents, and students, ESSP MAP creates online profiles of social environments and self-perceptions of students at the school, group, and individual levels. Online materials help schools interpret and use the profiles to establish tiered systems of supports (specific interventions) and identify students in need of those interventions.

A study examined the association between ESSP MAP use and student reading achievement for 10 cohorts of grade 3 students and 1 cohort of grade 4 students through their grade 5 year for one North Carolina district (Bowen, Thompson, & Powers, 2012). Teachers and school staff for two of these cohorts (grade 3 students in 3 schools and grade 4 students in another school, totaling approximately 830 students) were provided with ESSP MAP data and received training

⁷ Easton-Brooks and Davis (2007) examined effects for only those students from low-income families; thus, the sample included 187 Black students. Graves (2011) used a larger sample, including 2,500 Black students in grade 3.

on how to use the data. Eleven other cohorts of grade 3 students in six schools (approximately 7,900 students) served as a comparison group. The reading proficiency trajectories of the two cohorts of students whose teachers were trained to use ESSP MAP data were compared to the reading proficiency trajectories of the other 11 cohorts of students. The results show that Black students in the cohorts exposed to the ESSP MAP intervention experienced greater annual percentage-point gains in reading proficiency rates compared to Black students in the comparison cohorts (Bowen et al., 2012). Even though the study used a quasi-experimental design (QED), it was classified as providing promising evidence because it did not show that the groups were equivalent at baseline but controlled for selection factors statistically.

Good Behavior Game with enhanced academic curriculum

The Good Behavior Game with enhanced academic curriculum is a classroom-centered intervention that aims to address aggressive behavior and poor academic achievement (Barrish, Saunders, & Wolf, 1969). The Good Behavior Game involves assigning students into teams. To “win,” teams cannot exceed a certain threshold of disruptive behaviors. The intervention also involves curricular enhancements such as interactive read-alouds, journal writing activities, and directed thinking activities. Teachers receive 60 hours of training and direct supervision of classroom implementation.

The longitudinal associations between the Good Behavior Game/enhanced academic curriculum and student achievement in ELA and math and high school graduation were tested with a randomized controlled trial (RCT; Bradshaw, Zmuda, Kellam, & Ialongo, 2009). Twenty-seven first-grade classrooms were randomly assigned to one of three conditions: the Good Behavior Game/enhanced curriculum intervention, the Family School Partnership intervention, or a no-intervention comparison condition. The sample had 653 grade 1 students in 27 classrooms in nine schools in a single district. Black students made up 87 percent of the sample. All students were tracked through age 19. The results showed statistically significant positive associations between the Good Behavior Game/enhanced curriculum on students’ grade 12 test scores and high school graduation. This study did not meet WWC standards without reservations because insufficient information about student attrition was provided. Nor did it meet WWC standards with reservations because no information was provided about the groups of students at baseline. Thus, the associations demonstrated within the study cannot be considered causal, and the findings cannot be classified as providing strong or moderate evidence. The analyses did control for other selection factors, and, therefore, the study findings were classified as providing promising evidence (Bradshaw et al., 2009).

Mentoring

The Benjamin E. Mays (BEMI) Institute is a mentoring program for Black male students who are identified as at risk for performing poorly in school. BEMI participants attend classes with Black

male peers and are taught by Black male teachers. Each BEMI participant is expected to meet frequently with a local Black business or community leader (that is, a mentor) who introduces the participant to the benefits of his profession. Then, each week, one of these mentors is invited to discuss his area of expertise with the entire class. Students enroll in the program at the beginning of grade 7, and the intervention includes two years of daily, all-male classroom instruction; weekly meetings with mentors; and monthly meetings among parents, students, and mentors (Gordon, Iwamoto, Ward, Potts, & Boyd, 2009).

One study of this program examined BEMI participants' grade 8 scores in math and reading on the Connecticut Mastery Test. The analytic sample size included 61 students in one school, including 29 BEMI program students and 32 comparison students with similar backgrounds. The analysis controlled for students' math achievement scores from grade 6, and the results showed that BEMI students scored significantly higher than non-BEMI students on grade 8 math achievement tests (Gordon et al., 2009). Because the study showed a statistically significant positive association with math achievement after controlling for student's prior achievement, the evidence was classified as promising.

Parental involvement

Parental involvement at home. One analytic study used data from the ECLS-K to explore whether student, family, and school-related factors were associated with gaps in literacy achievement for 2,296 students in kindergarten and grade 1 (356 of whom were Black; Chatterji, 2006). The analyses included two measures of parental involvement: the average amount of time that parents spent reading to their child per week, and the extent to which parents assisted their child with schoolwork at home. The researcher used a statistical model that controlled for various student and school factors, including prior reading achievement. The results for Black students indicated a statistically significant positive association between the amount of time that parents read to their child and students' achievement in reading (Chatterji, 2006). This study used correlational methods while statistically controlling for selection factors. Accordingly, the evidence from the study was classified as promising.

Parental involvement at school. Researchers used data from the National Education Longitudinal Study (NELS:88) to identify factors that predict standardized test scores in math for Black students in grade 12 (Carpenter, Ramirez, & Severn, 2006). Data came from a representative sample of 15,618 students who were in grade 12 in 1992. Of these, 1,660 students were Black. The analysis included teachers' perceptions of each student's parental involvement. The results indicated a statistically significant positive association between the involvement of Black students' parents in the school and the students' math achievement in grade 12. A similar association was evident for White students (Carpenter et al., 2006). This correlational study

provides promising evidence of an association between parental involvement and Black students' achievement.

Another analysis of the NELS:88 data examined students' likelihood of dropping out of high school (Carpenter & Ramirez, 2007). The sample for this analysis included 17,613 students, 2,010 of whom were Black. Analyses controlled for student and school factors, including students' prior achievement test scores. The results of the analysis indicated a statistically significant negative association between the involvement of Black students' parents in the school and the likelihood of those students dropping out of high school. The results from this correlational analysis also provide promising evidence supporting an association between the involvement of Black students' parents in the school and the achievement of those students.

Positive Action: Social-emotional and character development program

Positive Action (PA) is a social-emotional and character development program based on self-concept and social-ecological theories. These theories suggest that positive thinking and action lead to fewer negative behaviors, higher academic motivation, and higher academic achievement. The PA core curriculum contains six units, taught four days per week in 15- to 20-minute lessons for grades K–6 and two days per week in 20-minute lessons for grades 7 and 8. The units focus on self-concept, positive actions focusing on oneself, positive social-emotional actions focusing on others, and continual self-improvement. The program also includes training and climate development for teachers, counselors, families, communities, and schools.

The association between PA and Black students' reading and math achievement was studied using a cluster RCT with multiple cohorts of students tracked for five years (Bavarian et al., 2013). Data from 1,170 students and 14 schools were included in the analyses. The association between PA and students' academic achievement was analyzed using a statistical model that controlled for students' test scores in prior years. PA was positively and significantly associated with standardized reading scores for Black male students. No positive and statistically significant associations were found for other ethnic groups or subject areas (Bavarian et al., 2013). This study was classified as providing promising evidence because of the high attrition rate and the lack of information about students' baseline equivalence.

Student Success Skills

The Student Success Skills (SSS) program involves professional development and coaching for school counselors to lead classroom lessons focused on cognitive, social, and self-management skills. School counselors are trained and coached to implement the program. Counselors lead eight 45-minute lessons in classrooms during the fall and another four refresher sessions in the late winter/spring. The sessions focus on teaching 16 tools related to cognitive, social, and self-management skills. At the beginning and end of each session, students discuss their progress on

their personal goals and share how they are applying their new skills to academics. Students receive additional practice and support if needed.

Two RCTs and two QEDs examined the impact of the program on student achievement, but none of these studies analyzed the achievement outcomes of Black and White students separately (Brigman & Campbell, 2003; Brigman, Webb, & Campbell, 2007; Campbell & Brigman, 2005; Webb, Brigman, & Campbell, 2005). The data from the four studies were reanalyzed to determine the association between student participation in SSS and math and reading scores for different racial/ethnic subgroups (Miranda, Webb, Brigman, & Peluso, 2007). The combined sample involved 1,123 students (279 of whom were Black) in grades 5, 6, 8, and 9 in two districts. Students' scores on the Florida Comprehensive Assessment Test in math and reading at baseline (the April before the study) and after the study was implemented were analyzed to determine program impact, and the analytic model controlled for students' baseline scores. The researchers found that students exposed to the SSS program outperformed their nonparticipating counterparts in ELA and math, and the improvements of Black students were similar to those of White students (Miranda et al., 2007). Even though this study was based on data from RCTs and QEDs, insufficient attrition and baseline information was provided for these primary studies (or the reanalysis) to meet WWC standards. The reanalysis was classified as providing promising evidence because it found statistically significant positive associations for Black students after controlling for students' baseline achievement.

Teacher practices in the classroom

Development of teacher-student relationships

A study was conducted to examine the associations between teacher–student relationships and student behavioral and academic achievement outcomes. Forty-four Black students in grades K–6 who were at risk of being referred to special education for behavior problems completed a survey instrument that assessed their relationship with their teacher in terms of their desire to be psychologically closer to their teacher (referred to as *psychological proximity seeking*) and the emotional quality of the relationship (Decker, Dona, & Christenson, 2007). The 25 teachers of these 44 students also recorded their perceived relationships with the students. The researchers developed analytic models that examined the associations between the quality of the teacher–student relationship and students' performance on two literacy assessments: Curriculum Based Measurement: Oral Reading Fluency (Deno, 1986) and Letter-Naming Fluency from the Dynamic Indicators of Basic Early Literacy Skills (Kaminski & Good, 2002). Analyses also controlled for other selection factors. None of the relationship variables was associated with students' oral reading fluency, but psychological proximity seeking was positively associated

with letter-naming fluency. This study used a correlational design with statistical controls, thus resulting in a classification as promising evidence.

Formative assessment

Formative assessments are used to monitor student learning to adapt classroom instruction (Sadler, 1998). One correlational study examined the association between students' responses to nine survey items about their teacher's use of formative assessment and the students' reading achievement. Data came from a nationally representative sample of 5,233 15-year-old students (635 of whom were Black) who participated in the Program for International Student Assessment (PISA; Li, 2016). The study found statistically significant positive associations between teachers' use of formative assessment and students' scores on the PISA reading test, even after controlling for student- and school-level socioeconomic status. The associations were stronger for Black students than for White or Hispanic students (Li, 2016). The evidence from this study was classified as promising because the study was correlational and controlled for selection bias.

Grade-specific instructional focus in math

Kindergarten: Instructional focus on telling time, and estimating quantities and coin values.

Researchers have explored national datasets to see whether certain instructional practices are associated with students' academic achievement. In one study, ECLS-K data from students from low-income families were used to determine whether the opportunity to learn math in school and teachers' self-reported emphasis on particular math topics was associated with these students' scores on the ECLS-K end-of-year assessment in math (Wang, 2010). The sample consisted of 1,272 students (643 Black students and 629 White students) and their teachers. Analytic models controlled for students' prekindergarten math skills, whether they attended full-day or half-day kindergarten, and other student factors. The findings showed that teachers' reported emphasis on accurately telling time and estimating quantities and coin values was associated with higher math achievement for both Black and White children from low-income families (Wang, 2010). The findings from this study were classified as promising evidence because the study used a correlational design with statistical controls.

Grade 4: Instructional focus on measurement and estimation. Another study examined survey and math assessment data from a representative sample of 13,511 grade 4 students who participated in the NAEP. The survey data included responses to questions about 20 instructional practices, and analyses examined whether any of those practices were associated with reductions in the Black-White achievement gap, after controlling for other student-related factors (Wenglinsky, 2004a). The findings revealed only one practice—teacher emphasis on measurement and estimation—was associated with a decrease in the achievement gap on the math portion of the NAEP (Wenglinsky, 2004a). Additional caution is warranted in interpreting

this statistical association because the study did not explicitly describe the separate trends that can constitute the achievement gap (that is, it could be a reduction in the achievement of White students, rather than improvement in achievement among Black students). Because this study used a correlational design with statistical controls for other student factors, the findings were classified as promising evidence.

High expectations

A study was conducted to determine whether Black students whose math teachers communicate high expectations show higher math achievement (Woolley, Strutchens, Gilbert, & Martin, 2010). The study sample included 933 Black students in grades 6–8 from 13 schools in seven districts. Students completed a survey that included items about their math teachers' communication of expectations. High expectations were defined as the aggregation of students' responses to five survey items about the degree to which their teacher expressed confidence that students can learn mathematical concepts. The students also completed the SAT-10 math assessment. The researchers' analyses controlled for student motivation, amount of studying, and math anxiety. The findings indicated that Black students whose teachers communicated high expectations had higher SAT-10 math scores (Woolley et al., 2010). This correlational study also serves as promising evidence for setting high expectations.

Homework

Evidence on the association between homework and Black students' achievement also comes from the previously mentioned analysis of NELS:88 data (see the description of the study supporting parental involvement at school; Carpenter et al., 2006). That same analysis included students' responses to survey questions on the amount of homework completed per week. A correlational analysis that controlled for various student characteristics examined the association between homework and grade 12 students' achievement on the NELS:88 math test. For Black and White students alike, the results indicated positive associations between time spent on homework and standardized test scores in math (Carpenter et al., 2006). The evidence supporting this association was classified as promising because it comes from a correlational study that controlled for selection factors statistically.

Instructional reform practice in math

The correlational study that provided evidence on teachers' high expectations also examined whether teachers' instructional reform practices in math were related to students' math achievement (Woolley et al., 2010). The sample was the same as that described previously for the homework intervention. The survey that students completed also included items about their math teachers' use of specific instructional reform practices (defined as practices requiring students to explain their reasoning and perform sustained work on challenging tasks). The

results indicated that teachers' use of reform practices was significantly associated with the SAT-10 math scores of Black students (Woolley et al., 2010). The findings from this correlational study were classified as providing promising evidence.

Time on task in math

Another study used data from the ECLS-K to determine whether the characteristics of teachers and teachers' instructional practices were associated with kindergarten and grade 1 students' performance on the ECLS-K math assessment (Desimone & Long, 2010). Analyses involved data from 10,980 kindergarten and grade 1 students, 661 of whom were Black. Analytic models controlled for the initial assignment of students to teachers and other student-level factors. The findings showed that time on task in math (teachers' reported amount of time conducting instruction in math) was significantly positively associated with math achievement for Black students in grade 1 but not associated with math achievement for White students. For kindergarten students, time on task was unrelated to math achievement for both Black and White students (Desimone & Long, 2010). This finding offers promising evidence that time on task for math is positively associated with the math performance of Black students in grade 1.

NAEP data from a nationally representative sample of 15,693 grade 8 students also were used to examine the association between math instructional practices and the Black-White achievement gap in math (Wenglinsky, 2004b). After controlling for other student and teacher characteristics, the study found that time on task in math was associated with a reduction in the Black-White achievement gap (Wenglinsky, 2004b). Similar to the study that was conducted using NAEP data for grade 4 students (Wenglinsky, 2004a), the results from this study did not show whether the associations with the Black-White achievement gap reflect positive associations for Black students, negative associations for White students, or both. Caution is therefore warranted when interpreting these findings. This study also used a correlational design with statistical controls; thus, the findings from this study were classified as promising evidence.

Supplemental interventions

The literature review also identified four interventions that are supported by promising evidence, yet are not implemented in any specific level of the K-12 education system. These interventions include those that are only partially implemented by school staff, those that occur outside the normal school hours, and those that can be implemented anywhere. These four interventions are summarized in the subsections that follow.

Out-of-school-time programs

Out-of-school-time (OST) programs are defined as programs that occur outside regular school hours, including extracurricular activities and after-school programs, clubs, sports, and lessons. One study was identified that examined the association between participation in OST programs

(that is, the number of hours participating in OST programs) and scores on college entrance examinations (that is, the ACT and SAT) for 2,363 Black scholarship applicants from single-parent homes (Nagle, 2013). On the scholarship form, students entered the total number of hours of OST participation from their junior year through the time they filled out the application (midway through their senior year), excluding summer activities. The average of students' scores on the ELA and math portions of the ACT and SAT served as the achievement measure. The analyses controlled for parents' education level and students' grade point average. The regression results showed that Black students' participation in OST programs was positively associated with their college entrance examination scores (Nagle, 2013). This study was classified as providing promising evidence because it was a correlational study with statistical controls for parents' education level and students' grade point average.

Summer reading program with free books

A summer reading program targeted students in grade 4 and had two components. First, teachers participated in two-hour training sessions on how to teach specific reading strategies and how to encourage students to read at home. Teachers then led these reading lessons with their students and provided encouragement during the last month of the school year. Second, students received eight books during July and August at no cost to them or their families.

In an experimental study, researchers examined whether students' participation in the intervention was associated with higher scores on reading portion of the Iowa Test of Basic Skills the following fall (Kim, 2006). The study was conducted with 552 grade 4 students from 10 high-poverty, multiracial schools in one district, and 93 of those students were Black. Students were randomly assigned to participate or not participate in the intervention during the summer. The researchers estimated the effect of the intervention for different racial/ethnic groups separately, with each statistical model controlling for students' reading achievement scores from the spring (before participation in the intervention). This study is likely to meet WWC standards without reservations, and for this particular sample, the data suggested that the intervention produced statistically significant improvements on Black students' achievement in reading. However, no effects were found for other racial/ethnic groups (Kim, 2006). The comparisons between Black students who were assigned to the summer reading program and Black students who were assigned to the comparison condition did not meet the sample size or multiple-settings requirements needed for a strong evidence classification. However, because the analytic models in the study controlled for selection bias, this study was classified as providing promising evidence.

Participation in urban debate leagues

Urban debate leagues (UDLs) promote and organize interscholastic debates through partnerships with urban public school systems. Most UDLs are for high school students

(grades 9–12), although some middle school leagues exist. In schools, debate is either embedded in an existing course or established as a distinct course. Through UDLs, universities are able to provide students with extra support, including mentoring opportunities. The skills needed to succeed in UDL competitions are aligned with the reading skills that students need at the secondary level: comprehension of complex texts, the ability to gather evidence from research, the ability to compare authors' claims, and the ability to distill key arguments from text (ACT, 2006; Mezuk, 2009). Thus, participation in UDLs may help Black students strengthen their reading skills and show stronger ELA achievement.

A study compared the ACT verbal scores (ELA achievement), graduation rates, and dropout rates of 458 Black male UDL participants in grades 9–12 with those of 2,156 matched students (Mezuk, 2009). Statistical models controlled for students' grade point averages and scores on the Illinois Standards Achievement Test in grade 8. The findings indicated that participation in a UDL was positively associated with the likelihood of scoring above the English and reading ACT benchmark scores and graduating from high school. Participation also was negatively associated with the likelihood of students dropping out of high school (Mezuk, 2009). This QED was classified as providing promising evidence because the two groups did not have equivalent achievement scores at baseline, but the baseline scores were statistically controlled.

Self-affirmation

Social identity threat is a social psychological theory that suggests that environmental cues that remind individuals that they are members of a marginalized group can lead those individuals to behave in a way that confirms the stereotypes of their group (Steele, Spencer, & Aronson, 2002). For example, according to the theory, when female students make up a minority in a math class, they may be reminded of the stereotype that females may not be as good in math as males, and they therefore perform accordingly. Social identity threat also may contribute to achievement gaps in schools in which White students outnumber Black students. The theory suggests that in such circumstances, Black students' awareness of their minority status within the school may remind them of the stereotype that Black students do not perform as well academically as White students do, thus perhaps prompting them to behave in accordance with that stereotype. An intervention that prompts Black students to think about their personally important values (that is, stimulates self-affirming thoughts) may help reduce achievement gaps caused by social identity threat (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009).

A study tested whether a self-affirmation intervention can improve Black students' academic performance, especially in schools in which White students outnumber Black students (Hanselman, Bruch, Gamoran, & Borman, 2014). Eleven middle schools in Madison, Wisconsin, participated in the study, and 910 students in grade 7 (including 161 Black students) were randomly assigned to write about values that were important to them (self-affirmation

condition) or write about values that were not important to them or about something procedural (comparison condition). In addition, the researchers distinguished between schools that presented a high-identity threat environment (that is, schools having a smaller ratio of Black to White and Asian students and a larger Black-White achievement gap during the previous school year) from those that presented a low-identity threat environment. The study found that in high-threat environments, Black students who responded to the self-affirmation writing prompts scored significantly higher on the language usage test from the Measures of Academic Progress, after controlling for other student and school characteristics. No statistically significant effects were found on Black students' math or reading achievement, nor were any self-affirmation effects found for White students. Although this study was an RCT and would likely meet WWC standards (assuming low attrition of students over the course of the study), it was conducted at only one site. The findings were therefore classified as promising evidence.

Interventions absent from this review

Three interventions were omitted from this review because the studies showing positive associations with favorable outcomes for Black students were negated by statistically significant countervailing findings. The three interventions were placement of students in advanced classes (Anderson, 2016; Byun, Irvin, & Bell, 2015), mixed-race learning environments (Goza & Ryabov, 2009; Grana, 2010), and hiring teachers of the same race as students (Banerjee, 2013; Egalite, Kisida, & Winters, 2015; Gershenson, Hart, Lindsay, & Papageorge, 2017). Additional studies need to be conducted to gain a better understanding of the conditions under which these interventions are beneficial or harmful.

Implications

For this literature review, interventions were defined as policies, practices, programs, or other malleable factor that may be associated with a student outcome. This literature review identified 24 studies providing promising evidence for 22 interventions. No interventions were supported by strong or moderate evidence.

The 22 interventions identified in this review need to be tested further

One implication for policymakers, educators, and researchers is that none of the interventions described in this report have yet demonstrated clear, unequivocal evidence of *causing* beneficial results for Black students. Until the effectiveness of the 22 interventions are tested with larger samples, more sites, and research designs that meet WWC standards, policymakers and educators who wish to improve the academic performance of Black students might benefit from taking a more cautious approach to adopting one or more of these interventions.

Some interventions need better articulation

For malleable factors identified through secondary analysis of large datasets (such as data from the Early Childhood Longitudinal Study and the National Education Longitudinal Study), the interventions may need to be better articulated in order to implement and evaluate them. For example, the promising evidence for “parental involvement” was based on parents’ survey responses and teachers’ survey responses that were part of the National Education Longitudinal Study dataset (Carpenter, Ramirez, & Severn, 2006). To better articulate a parental involvement intervention, a district may want to develop messages to parents on how to get involved with their child’s education and messages to teachers and school leaders on how to encourage parental participation. Districts also may want to identify parental participation activities that minimize potential risks to children and develop parental volunteer application and screening procedures. Improved articulation can also help with evaluating the intervention’s effects.

The intervention roll-out strategy can help accumulate additional evidence

Decisionmakers who are considering adopting one of the better articulated interventions as a way of improving the achievement of Black students may want to roll out implementation in a way that allows rigorous testing. By piloting interventions in a subgroup of districts or schools, decisionmakers can study the interventions’ effects with a more rigorous design. The most rigorous design—one capable of meeting WWC standards—would involve choosing districts or schools to implement the intervention through a random assignment process. For example, if state policymakers decide to implement the summer reading program with free books (see Kim, 2006), piloting the intervention with a randomly chosen group of eligible school districts would allow the policymakers to compare outcomes among Black students in districts implementing the intervention with those of Black students in districts that did not implement the intervention. Such a comparison could provide strong evidence, according to ESSA standards (U.S. Department of Education, 2016). A coordinated, comprehensive approach may be needed to improve the achievement of Black students relative to their White peers.

The 22 interventions identified in this review varied by level of the education system, by grade-level focus, and by content-area focus. Policymakers, district and school administrators, and teachers who want to reduce the gap in academic achievement between Black students and their White peers may attain better results by working together to develop a comprehensive approach. Stakeholders working on the issue may want to adopt multiple interventions (some at different levels of the education system, different grade levels, and different content areas). Yet implementing multiple interventions in an uncoordinated way may produce policy incoherence and other unintended consequences (Honig & Hatch, 2004). A multidisciplinary working group or research-practice partnership—such as MAGRA, which prompted this research review—may help provide the necessary coordination.

This research review did not examine studies on prekindergarten-aged children, but research shows that disparities in academic achievement between Black children and their White peers are evident even before they enter the K–12 education system (Jacobson Chernoff et al., 2007; Reardon & Portilla, 2016). Accordingly, stakeholder groups working on a coordinated approach to reducing achievement gaps between Black and White students may find it beneficial to include stakeholders who work with younger-aged children and their families.

Limitations of the research review

This research review has four limitations.

First, the purpose of the review was to identify interventions associated with Black students' achievement for which strong, moderate, or promising evidence exists. However, identifying a single study supportive of an intervention does not guarantee that the intervention will show similar associations in all settings, nor does it mean that other studies on the same intervention found similar associations. In accordance with the descriptions of the ESSA tiers, the literature review team purposely omitted interventions for which significant associations with favorable and unfavorable outcomes were found (as mentioned in the previous section). However, the literature review team did not adjust the classification of evidence to account for studies of the same interventions that show nonsignificant associations with Black students' achievement outcomes.⁸ If the associations between interventions and outcomes (both significant and nonsignificant) were viewed collectively, one might conclude that an intervention holds less promise than its ESSA tier label would suggest.⁹

Second, the list of interventions identified in this report is not exhaustive. Terms related to Black students were included in the literature database searches. Thus, only studies that explicitly mention Black students in the abstract, keywords, or descriptors could be identified. However, there likely are studies of other interventions that included separate analyses of Black and White subgroups but neglected to mention racial differences in the abstract, keywords, or descriptors. Expanding the literature search to identify such studies was not feasible given time and budget constraints.

Third, there could be interventions associated with improved achievement among Black students, but were not included in this review because less than 80 percent of the sample was

⁸ Interventions for which nonsignificant associations were found are listed in table C1 in appendix C.

⁹ The U.S. Department of Education's nonregulatory guidance on ESSA recommends that "SEAs, LEAs, and other stakeholders should consider the entire body of relevant evidence" (U.S. Department of Education, 2016, p. 8).

made up of Black students or because the researchers did not perform subgroup analyses. Such studies would not have passed through the full-text screening process.

Fourth, the classification of studies using ESSA tiers of evidence fails to consider the magnitude of statistically significant effects, the degree to which studies had samples (or subsamples) of Black students large enough to detect effects, and the magnitude of effects from studies that found no statistically significant associations. Specifically, some interventions may show statistical associations with Black student outcomes, but the size of the association can be quite small. Similarly, some studies may show large associations with Black student outcomes, but because of a small sample size, they fail to reach the level needed to be statistically significant. The magnitude of effects for some interventions may be found in WWC reports. To determine the magnitude of associations for other interventions, readers may need to consult the research reports cited in this review or consult a local evaluator.¹⁰

¹⁰ The magnitude of effects for interventions may be found in the studies or in WWC intervention reports. Resources for calculating and interpreting effect sizes can be found in books on meta-analysis (for example, Cooper, 2010) or the What Works Clearinghouse Procedures Handbook (version 4; What Works Clearinghouse, 2017). However, these resources are targeted toward researcher audiences and may be too technical for policymakers and practitioners.

Appendix A. Literature review methodology

This appendix describes the processes that the Regional Educational Laboratory Midwest project team followed while conducting this literature review. The processes for searching for relevant studies, screening studies, recording study features, and classifying the findings from eligible studies according to the ESSA evidence tiers are described in detail.

Eligibility requirements for studies

To be eligible for this research review, studies had to meet the following requirements:

- **Year of publication:** Studies published in 2002 or later were eligible. This 16-year period was recommended by members of the Midwest Achievement Gap Research Alliance (MAGRA) and corresponds to the years in which state and local education agencies have worked to reduce Black-White achievement gaps (as mandated by the No Child Left Behind Act of 2001).
- **Types of publications:** The research review team sought out quantitative studies. Such studies most likely would be published as doctoral dissertations, articles in research or practitioner journals, or reports/working papers from research organizations and cited in research literature databases.
- **Study sample:** Studies had to involve students in grades K–12 who reside in the United States, its territories, or tribal entities. The sample either had to be more than 80 percent Black (a criterion similar to that set by Bandy & Moore, 2011) or include a distinctly Black subgroup.
- **Study design:** Only studies that produced evidence in tiers I, II, or III of ESSA were included in the review. Such studies had to employ random assignment of units to experimental conditions (that is, a randomized controlled trial or RCT), a quasi-experimental design (QED), or a correlational design with statistical controls for selection biases.
 - *Randomized controlled trial:* Per the U.S. Department of Education’s nonregulatory guidance on interpreting ESSA (U.S. Department of Education, 2016), for study results to be classified as strong evidence (tier I), the study must meet What Works Clearinghouse (WWC) standards without reservations (that is, a study must involve

- the random assignment of units to experimental conditions, otherwise known as an RCT) and have low attrition.¹¹
- *Quasi-experimental designs*: To be classified as moderate evidence, a study had to be a QED and meet WWC standards with reservations. Those standards require that the groups being compared be equivalent on important measures prior to the study (at the baseline). Groups are considered equivalent if they differ by 0.05 standard deviation or less on baseline measures or if they differ by 0.25 standard deviation with baseline measures used as covariates in analytic models (What Works Clearinghouse, 2014).
 - *Correlational studies with controls*: The review team considered studies that used partial correlation analyses, analyses of covariance, or regression analyses with covariates. To be eligible for the research review, the studies needed to account for other factors that could influence the relationship between an intervention and an outcome (that is, using covariates or control variables). Relevant covariates include past achievement and demographic variables, such as family socioeconomic status.
 - **Interventions**: Eligible studies had to examine an intervention, defined as a policy, a practice, a program, or some other malleable factor. Malleable factors are those that can be changed or altered through changes in education policy, practices, or programs.¹²
 - **Outcomes**: To be eligible, studies had to show statistically significant associations between an intervention and students' standardized test scores in English language arts (ELA, including reading and literacy) or math or high school graduation or dropout rates.

Search for literature

To find studies that met the eligibility criteria, the review team searched the following databases or websites:

- Education Resources Information Center (ERIC) online platform
- Databases accessible through the EBSCO platform:
 - Academic Search Premier

¹¹ The WWC has two sets of boundaries for determining if a study lost too many research participants (either clusters or students) between randomization and post-test to consider the findings trustworthy. The two sets of thresholds are the “conservative boundary” (requires more units to stay in the study for the study to be valid) and the “liberal boundary” (allows more loss of participants before questioning the validity of the study; What Works Clearinghouse, 2014, pp. 12–13). The WWC-certified reviewer who rated whether studies in this review met WWC standards used the liberal boundary.

¹² This definition is consistent with that used by the Institute of Education Sciences (see Earle et al., 2013).

- Education Research Complete
 - Education Source
 - ERIC (EBSCO-based)
 - Library, Information Science & Technology Abstracts
 - Mental Measurements Yearbook with Tests in Print
 - PsycINFO
 - Psychology and Behavioral Sciences Collection
 - SocINDEX with Full Text
- Google Scholar

MAGRA members also recommended references for the study team to examine. Those reports were subjected to the same screening process as references found in online databases.

The review team developed parallel search strategies and keywords for each search platform. Keyword searches included three phrases: [an achievement outcome] AND [an age group/grade-level band] AND [indicator of the achievement gap topic or target student population]. The team used both traditional keyword strings and database-specific identifiers (that is, predefined search terms). The team entered the traditional keyword strings into the thesaurus of each major database to identify database-specific identifiers. The keyword strings and search identifiers (table A1) were systematically entered into the databases based on each database's search requirements.

The online ERIC platform uses simple search phrases and descriptors rather than complex Boolean logic. The native ERIC platform, accessed via the online webpage, includes recent entries that may not yet be available through the EBSCO platform. The native platform also suggests descriptors by education level that produce the greatest number of results. The review team used ERIC's index of research reports from various websites to search for reports from relevant foundations, associations, and organizations. The team searched the ERIC database using both the native platform and the multi-database EBSCO platform to maximize the results.

Within the EBSCO platform, the review team conducted searches using both traditional keyword strings and descriptors.

In Google Scholar, the review team screened the first 100 results per search string, which represented the best matches for that search string and the most that could be screened given available resources. Each string included an outcome and age/grade band along with the indicators for the target student population (see table A1).

Table A1. Keywords and descriptors used in the literature search

Platform, type of search, and topic	Keywords or descriptors used in search
EBSCO general keywords: Outcomes	Mathematics assessment; Math assessment; Mathematics achievement test, Math achievement test, Reading assessment, English language assessment; English language arts assessment; Language achievement test, ELA assessment; ELA achievement test; Standardized test; High school graduation; High school graduate; Finish high school; High school completion; Drop-out; Dropout; Achievement gap; Opportunity gap
EBSCO general keywords: Grade levels	Elementary school; Elementary education; Primary school; Primary education; Grade school; Kindergarten; Grade 1; Grade 2; Grade 3; Grade 4; Grade 5; 1st grade; 2nd grade; 3rd grade; 4th grade; 5th grade; 6th grade; Junior high; Middle school; 7th grade; 8th grade; Grade 7; Grade 8; High school; Secondary school; Secondary education; 9th grade; 10th grade; 11th grade; 12th grade; Grade 9; Grade 10; Grade 11; Grade 12; Freshm*; Sophomores; Juniors; Seniors
EBSCO general keywords: Target population	African American; Black; African-American
EBSCO specific descriptors: Topic or target population	DE "Achievement gap" DE "African American students" DE "Black students" DE "Black high school students" DE "Black students – Research" DE "African Americans" DE "Black youth – Education" DE "African Americans – Education" DE "African American children" DE "African American children – Education" DE "African Americans" DE "Black students" DE "African American achievement" DE "Blacks" DE "Black Students – Psychology" DE "African Americans" DE "African American high school students" DE: "Black high school students"
ERIC-specific descriptors: Achievement outcomes	Descriptor: "achievement gap" Descriptor: "mathematics achievement" Descriptor: "reading achievement" Descriptor: "graduation rates"
ERIC-specific descriptors: Grade bands	Descriptor: "middle schools" Descriptor: "elementary education" Descriptor: "high schools" Descriptor: "secondary education" Descriptor: "elementary secondary education"
ERIC-specific descriptors: Target population	Descriptor: "African American students"

Platform, type of search, and topic	Keywords or descriptors used in search
Google Scholar-specific keyword strings	Elementary school math assessment Black students Elementary school math assessment African American students Elementary school ELA assessment African American Black students Elementary school Reading assessment African American Black students Middle school math assessment African American Black students Middle school ELA assessment African American Black students Middle school reading assessment African American Black students High school math assessment African American Black students High school ELA assessment African American Black students High school reading assessment African American Black students High school graduation rates African American Black students High school dropout rates African American Black students Elementary school achievement gap African American Black students Middle school achievement gap African American Black students High school achievement gap African American Black students

* Signifies “any word beginning with”; DE is the syntax used by EBSCO to identify descriptor terms (rather than keywords). ELA is English language arts.

Source: Authors' compilation.

The results of each database search produced a reference for each identified study and the study abstract (or brief summary). The files containing the search results from the various databases were compiled into a single file, and this larger file was uploaded into EndNote software. The project team used EndNote to remove duplicate references (or references that were identified in multiple databases). This deduped file was then entered into a customized screening and coding database. The customized database also included an algorithm that identifies duplicate references by looking for matching field names (specifically, author[s], title, and year). Any remaining duplicates were removed using this algorithm.

Screening process and study eligibility criteria

The principal investigator developed a screening and reviewing protocol that described the methodology used for this review, the criteria for inclusion in the review, definitions of key terms and examples of hard-to-screen abstracts and report summaries, and the judgments that should be made in such instances. Once this protocol was completed, the principal investigator trained four other staff to serve as screeners.

Screener training

Screener training consisted of two, one-hour training sessions, followed by weekly calibration meetings. During the first training session, the principal investigator introduced the screeners to the review topic and protocol and the process in which screening judgments were to be made. The principal investigator then led the screeners through an exercise in which they all read five study abstracts and shared how they would answer seven screening questions for each abstract. For each question, screeners could answer “Yes,” “No,” or “Do not know.”

1. Was the study report a published journal article, technical report, or dissertation?
2. Was the study report published between 2002 and 2017?
3. Was the study report published in English?
4. Does the study report include at least one study with students in grades K–12 (approximately ages 5–18) in the United States?
5. Does the study report include at least one study that examines one of the following educational outcomes:
 - a. Math or ELA achievement (for example, standardized test, educational assessment, math or language arts examination)?
 - b. High school completion (for example, graduation rates, dropout rates)?
 - c. Achievement outcomes not otherwise specified?
6. Was the intervention (educational policy, program, practice, malleable factor) described in relation to educational outcomes for Black/African-American students or all students?
7. Did the study use a quantitative research design?

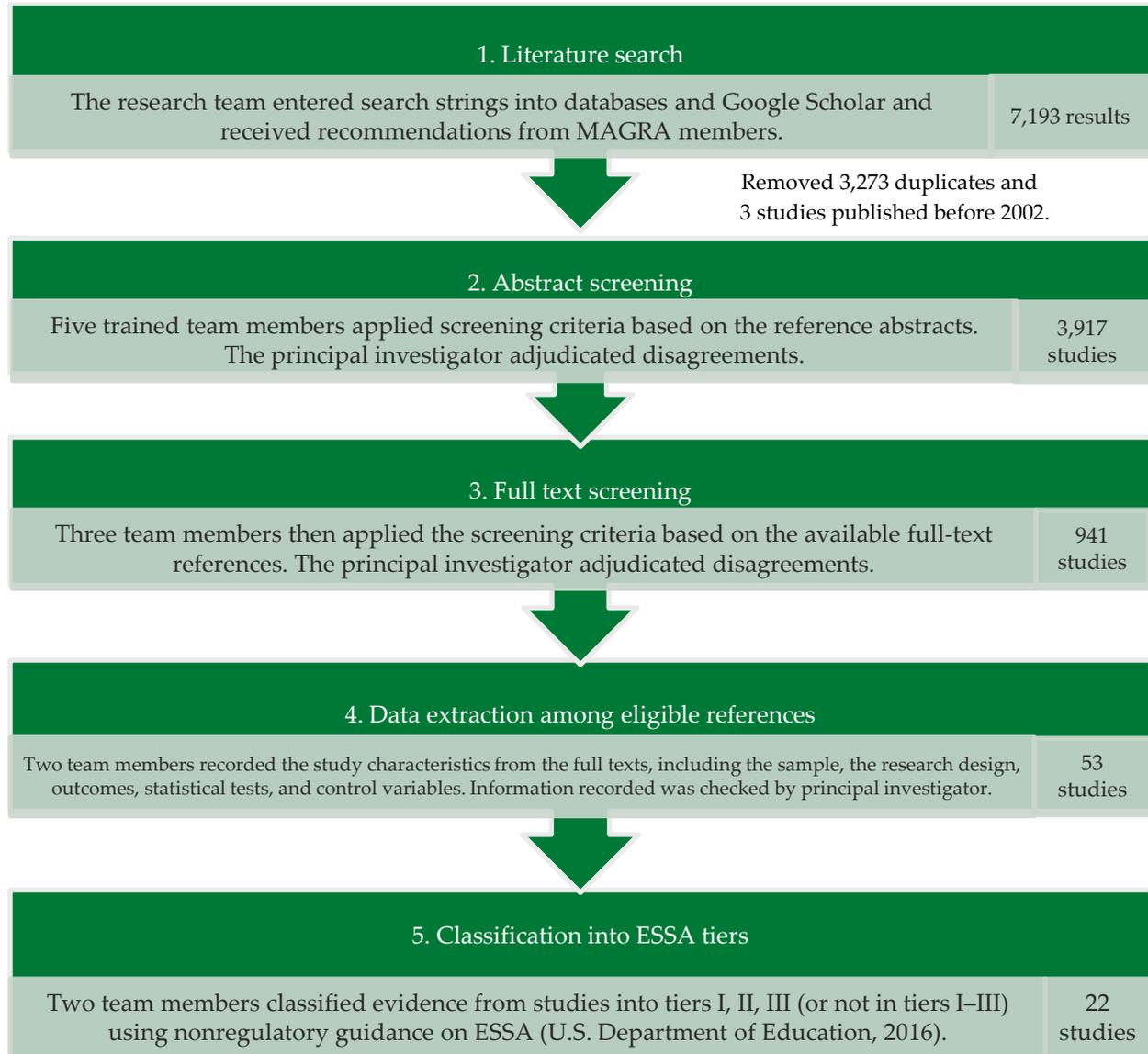
The team discussed differences of opinion on how each screening question would be answered for each study. The discussion continued until all screeners expressed an understanding of how to apply the eligibility criteria.

Between the first session and the second session, screeners were asked to read 10 more study abstracts and answer the screening questions for each abstract to the best of their ability. Prior to the second training session, the principal investigator examined screeners' judgments for accuracy and consistency. The second training session focused on discussing discrepancies between screeners' answers and the principal investigator's answers. Once screeners expressed an understanding of the screening questions and how to make judgments, they each were assigned to review 20 percent of the abstracts, with two team members screening 10 percent of the abstracts (the principal investigator screened abstracts as well). Each week, the principal investigator checked the ratings for these shared abstracts and held a conference call in which discrepant responses were discussed and adjudicated.

Screening process

Studies identified through the systematic searches went through two rounds of screening to determine whether they were eligible for inclusion in the review. The first screening round occurred by reviewing the abstract, and the second screening round occurred by reviewing the full text (figure A1). During both rounds, screeners answered the seven previously listed screening questions.

Figure A1. The interventions with strong, moderate, or promising evidence were identified through a research review involving five steps



Source: Authors' compilation.

Screening results

During the abstract screening round, five team members screened 3,917 report abstracts and judged 2,976 to be ineligible, leaving 941 eligible for full-text review. The project's research librarian was unable to locate 234 of these reports, and 146 other reports were located but could not be obtained within the time available for the review. Of the latter, 116 were dissertations, 21 were books, and 9 were journal articles. Ultimately, three team members screened 561 full-text

reports. The screeners found that 508 reports did not meet the previously listed criteria for inclusion. The 53 remaining reports progressed to the initial ESSA classification phase.

Throughout the screening phase, two screeners independently screened 10 percent of the abstracts and full-text reports. During abstract screening, the agreement rate was 89 percent across all judgments, and 12 percent of those inconsistencies (or 1.3 percent of the total) resulted in misclassification of a report as potentially eligible or ineligible. The principal investigator resolved any disagreements. For the full-text screenings, the agreement rate was 77 percent. The majority of those disagreements (87 percent) reflected instances in which one screener marked “Do not know” to an eligibility question whereas the other screener marked either “Yes” or “No.” The principal investigator checked the ratings for all studies for which a screener responded with “Do not know”. Only 2 percent of the disagreements resulted in a misclassification of a study as eligible or ineligible, and the principal investigator corrected these misclassifications.

Data extraction

Two team members then examined the 53 study reports that passed through the screening. These team members recorded the key features for each study. These key features included the full reference, the research design, the characteristics of the sample, the statistically significant associations found, operational definitions of outcomes, statistical tests performed and control variables used in analytic models. These codes were checked during the ESSA classification stage.

Classification of findings according to ESSA tiers of evidence

During the ESSA classification phase, two team members, one of whom was a WWC-certified reviewer, examined the 53 study reports and classified the findings from each study as strong, moderate, or promising based on ESSA’s tiers of evidence (U.S. Department of Education, 2016; table A2) and WWC standards (version 3.0; What Works Clearinghouse, 2014). RCT and QED studies were examined for sample attrition information and information on equivalence of groups at baseline. Studies that did not meet the WWC’s attrition standard (liberal boundary) and baseline equivalence standards still could provide promising evidence (tier III) if the analyses controlled for other selection factors (such as, baseline achievement, or socioeconomic status). The study classification process resulted in the removal of 31 studies because they did not provide evidence that could be classified as strong, moderate, or promising or because their positive effects were negated by unfavorable effects in other studies. As a result, 22 study reports remained.

Table A2. Tiers of evidence from the Every Student Succeeds Act

Tier	Finding	Design	Sample	Other
Tier I (strong evidence)	Statistically significant association between intervention and favorable outcome	Meets WWC standards without reservations (can detect causal associations)	2+ sites AND 350+ students or 50+ schools, classrooms, teachers	No other studies meeting WWC standards that show statistically significant association with unfavorable outcome
Tier II (moderate evidence)	Statistically significant association between intervention and favorable outcome	Meets WWC standards with reservations (can detect causal associations)	2+ sites OR 350+ students or 50+ schools, classrooms, teachers	No other studies meeting WWC standards that show statistically significant association with unfavorable outcome
Tier III (promising evidence)	Statistically significant association between intervention and favorable outcome	Correlational study with statistical control for selection factors (cannot detect causal associations)	No sample size criteria	No other studies meeting WWC standards that show statistically significant association with unfavorable outcome

Note: WWC is What Works Clearinghouse.

Source: Based on U.S. Department of Education (2016).

Appendix B. Interventions supported by promising evidence

This appendix provides more information on the findings from the research review. Interventions that were found to be associated with Black students' achievement based on promising evidence are listed according to the level of education system at which they are implemented (table B1).

Table B1. Interventions associated with Black student academic outcomes and characteristics of studies that provide promising evidence for those associations

Level of education system	Intervention and its component processes	Student grade levels with favorable associations	Academic outcome	Study citation
State level	District Assistance and Intervention Teams: Financial resources and district support services led by trained and experienced consultants	Grades 2–11	Math	Strunk & McEachin (2014)
District level	Hiring certified teachers: hiring teachers with standard certificates and assigning them to teach classes for which they are certified	Kindergarten, grade 3	ELA	Easton-Brooks & Davis (2009); Graves (2011)
School level	Elementary School Success Profile Model of Assessment and Prevention: Tool for developing profiles at the school, group, and student levels using data from teachers, parents, and students	Grades 3–5	ELA	Bowen et al. (2012)
	Good Behavior Game with enhanced academic curriculum: Competition for good behavior; interactive read-alouds, journal writing, directed thinking; teacher training and supervision	Grades 1–12	ELA and math, graduation	Bradshaw et al. (2009)
	Benjamin E. Mays Institute (mentoring for Black male students): assignment to male Black teacher, mentoring relationship with Black community leader; presentations by community leaders	Grade 7, 8	Math	Gordon et al. (2009)
	Parental involvement at home: Parents spending time reading to their child	Grade 1	ELA	Chatterji (2006)
	Parental involvement at school: Parents' participation in school activities, communication with teachers about their children.	Grade 12	Math and dropout	Carpenter et al. (2006); Carpenter & Ramirez (2007)
	Positive Action: social-emotional and character-development curriculum with lessons on self-concept, positive actions focusing on oneself, positive actions toward others, continual self-improvement; also school climate development activities	Grades K–8	ELA	Bavarian et al. (2013)
	Student Success Skills: counselor-led sessions focused on cognitive, social, and self-management skills	Grades 5, 6, 8, 9	ELA and math	Miranda et al. (2007)
Classroom level	Development of student-teacher relationships: teachers' efforts toward becoming closer to their students, developing sense of psychological proximity with students	Grades K–6	ELA	Decker et al. (2007)
	Formative assessment: using data for purposes of providing feedback to students and adapting classroom instruction	Grades 9–10	ELA	Li (2016)

Level of education system	Intervention and its component processes	Student grade levels with favorable associations	Academic outcome	Study citation
Classroom level (continued)	Reform practices in math: asking students to explain reasoning, encouraging students to sustain work on challenging tasks	Grades 6–8	Math	Woolley et al. (2010)
	Grade-specific instructional focus in math: Math subjects emphasized by teachers for different grade levels; for kindergarten, emphasis is on telling time, estimating quantities and coin values; for grade 4, emphasis is on measurement and estimation	Kindergarten, Grade 4	Math; Black-White gap in math	Wang (2010); Wenglinsky (2004a)
	High expectations in math: communicating positive expectations about students' abilities to do math	Grades 6–8	Math	Woolley et al. (2010)
	Homework: students' performing school work outside school	Grade 12	Math	Carpenter et al. (2006)
	Time on task in math: amount of time that teachers spend providing instruction to students in math or leading math activities	Grades 1 and 8	Math; Black-White gap in math	Desimone & Long (2010); Wenglinsky (2004b)
Other	Out-of-school-time programs: student participation in activities/programs outside regular school hours, such as after-school programs, clubs, sports, and lessons	Grades 11–12	ELA and math	Nagle (2013)
	Self-affirmation in high-stereotype threat conditions: stimulation of self-affirming thoughts among Black students through having them think about personally important values	Grade 7	ELA	Hanselman et al. (2014)
	Summer reading program with free books: teacher-led instruction on reading strategies and encouragement to read at end of school year plus provision of eight free books to students in July and August	Grade 4	ELA	Kim (2006)
	Urban debate leagues: instruction on research and debating; participation in cross-school debate competitions	Grades 9–12	ELA, dropout	Mezuk (2009)

Note: ELA is English language arts; graduation and dropout may represent outcomes at the individual level (a student's choice to graduate or drop out of high school) or graduation and dropout rates, which are these outcomes at a broader level of aggregation.

Source: Authors' analyses.

Appendix C. Interventions with no associations or negative associations with Black students' outcomes

This appendix shows interventions for which tested associations were not statistically significant or statistically significant with an unfavorable outcome (table C1).

Table C1. Interventions showing no associations with Black students' achievement outcomes or significant associations with unfavorable outcomes

Level of education system	Intervention	Black student grade level and specific student subgroups	Outcomes showing no associations with the intervention	Outcomes showing unfavorable associations with the intervention	Study citation
State	High school exit examination with no alternate pathway for graduation	Grade 12		Dropout	Hemelt & Marcotte (2013)
	High school exit examination with alternate pathway for graduation	Grade 12	Dropout (female students)	Dropout (male students)	Hemelt & Marcotte (2013)
	West Virginia's professional development on closing the achievement gap	Grade 4	ELA and math		White, Hixon, Hammer, Smith & D'Brot (2010)
District level	Full-day kindergarten	Grades K–1	Math		Desimone & Long (2010)
	Hiring and placement of teachers based on <ul style="list-style-type: none"> • Amount of experience • Amount of education • Certification • Amount of professional development 	Grades K–1	Math		Desimone & Long (2010)
	Hiring and placement of fully certified teachers	Grade 1	ELA		Chatterji (2006)
	Hiring and placement of teachers with more education	Grades K–3	ELA		Graves (2011)
	Hiring and placement of teachers with more content preparation and experience	Grade 11 male students who were high achieving in math in grade 9	Math		Anderson (2016)
	Hiring of Black teachers and student-teacher ethnic matching	Grades K–3	ELA and math	ELA and math	Banerjee (2013)
	Hiring of Black teachers and assignment of all students to at least one Black teacher	Grades K–5	Math		Eddy & Easton-Brooks (2011)

Level of education system	Intervention	Black student grade level and specific student subgroups	Outcomes showing no associations with the intervention	Outcomes showing unfavorable associations with the intervention	Study citation
District level (continued)	School inputs <ul style="list-style-type: none"> • Mean class size • Mean school size • Percentage of students with individualized education plans 	Grade 1	ELA		Chatterji (2006)
	School racial composition	Grade 7	ELA and math		Hanselman et al. (2014)
	Homogeneous Black population (low-stereotype threat environment)	Grade 8	Math		Arbuthnot (2009)
School level	Placement of students into low-ability groups	Grades 1–3		ELA	Lleras & Rangel (2009)
	Placement of students in advanced math courses	Grade 12		Math	Byun et al. (2015)
	Teacher support in school	Grade 7	Math		Gutman, Sameroff, & Eccles (2002)
	Anxiety intervention program	Grades 3–5	ELA and math		Cooley-Strickland, Griffin, Darney, Otte, & Ko (2011)
	Dropout program	Grades 8–12	Dropout		Carpenter & Ramirez (2007)
	ESL program	Grades 8–12		Dropout	Carpenter & Ramirez (2007)
	Suspension	Grades 8–12		Dropout	Carpenter & Ramirez (2007)
	Retention in grade	Grades 8–12		Dropout	Carpenter & Ramirez (2007)
	Parent school involvement	Kindergarten	ELA		Graves & Brown Wright (2011)
	Parent school involvement	Grade 7	Math		Gutman et al. (2002)
Parent school involvement	Grade 12	ELA and math		Jeynes (2005)	

Level of education system	Intervention	Black student grade level and specific student subgroups	Outcomes showing no associations with the intervention	Outcomes showing unfavorable associations with the intervention	Study citation
Classroom level	Teacher-reported student–teacher relationship Student-reported emotional quality	Grades K–6 (students who are at risk for referral to special education)	ELA		Decker et al. (2007)
	Grade-specific instructional emphasis in math <ul style="list-style-type: none"> • Use of manipulatives • Use of worksheets and textbooks • Use of explanations and real-life math • Reading graphs and doing simple data collection 	Kindergarten	Math	Math (only for use of manipulatives)	Wang (2010)
	Grade-specific instructional emphasis in math: <ul style="list-style-type: none"> • Basic procedural topics • Advanced procedural topics • Conceptual topics 	Grades K–1	Math		Desimone & Long (2010)
	Grade-specific instructional emphasis in math: 19 instructional practices	Grade 4	Black-White math achievement gap		Wenglinsky (2004a)
	Grade-specific instructional emphasis in math 13 instructional practices	Grade 8	Black-White math achievement gap		Wenglinsky (2004b)
	Class time dedicated to math and reading instruction	Grade 1	ELA		Chatterji (2006)
	Time spent on reading	Grades K–3	ELA		Graves (2011)
	Flipped instruction	Grade 8	Math		Martin (2015)
	Differentiated learning context: <ul style="list-style-type: none"> • High communal learning • High self-regulated • Combination of high self-regulated and high communal context) 	Grades 3 and 5	Math		Jackson (2012)
	Other	Summer math/science programs in partnership with higher education institutions	Grade 11 (Males who were high achieving in math in grade 9)	Math	
After-school math/science program with mentors		Grade 11 (Males who were high achieving in math in grade 9)	Math		Anderson (2016)
Self-affirmation intervention alone		Grade 7	ELA and math		Hanselman et al. (2014)

ELA is English language arts. ESL is English as a second language.

Source: Authors' analyses.

References¹³

ACT (2006). *Reading between the lines: What the ACT reveals about college readiness in reading*. Iowa City, IA: ACT.

*Anderson, K. A. (2016). Examining organizational practices that predict persistence among high-achieving Black males in high school. *Teachers College Record*, 118(6), 1–26.
<https://eric.ed.gov/?id=EJ1100382>

Annie E. Casey Foundation. (2017). *Race for results: Building a path to opportunity for all children. 2017 Policy Report, Kids Count*. Baltimore, MD: Author. Retrieved December 15, 2017 from <http://www.aecf.org/m/resourcedoc/aecf-2017raceforresults-2017.pdf>.

*Arbuthnot, K. (2009). The effects of stereotype threat on standardized mathematics test performance and cognitive processing. *Harvard Educational Review*, 79(3), 448–472.
<https://eric.ed.gov/?id=EJ861868>

Bandy, T., & Moore, K. A. (2011). *What works for African American children and adolescents: Lessons from experimental evaluations of programs and interventions* (Fact Sheet Publication #2011-04). Washington, DC: Child Trends. <http://eric.ed.gov/?id=ED521831>

*Banerjee, N. (2013). *Effects of teacher diversity and student-teacher racial/ethnic matching in elementary schools on educational outcomes*. (Doctoral dissertation, University of North Carolina, Charlotte, 2013). <https://eric.ed.gov/?id=ED560382>

Barnett, S., Carolan, M., & Johns, D. (2013). *Equity and excellence: African-American children's access to quality preschool*. Center on Enhancing Early Learning Outcomes.
<https://eric.ed.gov/?id=ED555720>

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good Behavior Game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis*, 2(2), 119–124.

Barton, P. E., & Coley, R. J. (2010). *The Black-White achievement gap: When progress stopped*. Princeton, NJ: Educational Testing Service. <https://eric.ed.gov/?id=ED511548>

*Bavarian, N., Lewis, K. M., DuBois, D. L., Acock, A., Vuchinich, S., Silverthorn, N., et al. (2013). Using social-emotional and character development to improve academic outcomes: A matched-pair, cluster-randomized controlled trial in low-income, urban schools. *Journal of School Health*, 83(11), 771–779. <https://eric.ed.gov/?id=EJ1026798>

¹³ References preceded by an asterisk (*) are sources that met screening criteria and were evaluated for evidence. References preceded by a plus symbol (+) are sources recommended by the alliance.

- Bower, K. M., Thorpe, R. J., Rohde, C., & Gaskin, D. J. (2014). The intersection of neighborhood racial segregation, poverty, and urbanicity and its impact on food store availability in the United States. *Preventive Medicine, 58*, 33–39.
- *Bowen, N. K., Thompson, A. M., & Powers, J. D. (2012). A quasi-experimental test of the Elementary School Success Profile Model of Assessment and Prevention. *Journal of the Society for Social Work and Research, 3*(3), 178–196.
- *Bradshaw, C. P., Zmuda, J. H., Kellam, S. G., & Ialongo, N. S. (2009). Longitudinal impact of two universal preventive interventions in first grade on educational outcomes in high school. *Journal of Educational Psychology, 101*(4), 926–937. <https://eric.ed.gov/?id=EJ860910>
- *Brigman, G., & Campbell, C. (2003). Helping students improve academic achievement and school success behavior. *Professional School Counseling, 7*, 91–98. <https://eric.ed.gov/?id=EJ771103>
- *Brigman, G., Webb, L., & Campbell, C. (2007). Building skills for school success: Improving the academic and social competence of students. *Professional School Counseling, 10*, 279–288. <https://eric.ed.gov/?id=EJ767376>
- *Byun, S.-Y., Irvin, M. J., & Bell, B. A. (2015). Advanced math course taking: Effects on math achievement and college enrollment. *Journal of Experimental Education, 83*(4), 439–468. <https://eric.ed.gov/?id=EJ1071098>
- *Campbell, C., & Brigman, G. (2005). Closing the achievement gap: A structured approach in group counseling. *Journal for Specialists in Group Work, 30*, 67–82.
- *Carpenter, D. M., Ramirez, A., & Severn, L. (2006). “Gap” or “gaps”: Challenging the singular definition of the achievement gap. *Education and Urban Society, 39*(1), 113–127. <https://eric.ed.gov/?id=EJ745027>
- *Carpenter, D. M., & Ramirez, A. (2007). More than one gap: Dropout rate gaps between and among Black, Hispanic, and White students. *Journal of Advanced Academics, 19*(1), 32–64. <https://eric.ed.gov/?id=EJ786605>
- Carter, P. L., & Welner, K. G. (Eds.). (2013). *Closing the opportunity gap*. New York, NY: Oxford University Press.
- *Chatterji, M. (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. <https://eric.ed.gov/?id=EJ742196>

- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science* 324(5925), 400–403.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D. et al. (1966). *Equality of educational opportunity*. Washington DC: U.S. Government Printing Office. <https://eric.ed.gov/?id=ED012275>
- Collins, M. B., Munoz, I., & JaJa, J. (2016). Linking ‘toxic outliers’ to environmental justice communities. *Environmental Research Letters*, 11, 1–9. Retrieved December 8, 2017, from <http://iopscience.iop.org/article/10.1088/1748-9326/11/1/015004/pdf>
- *Cooley-Strickland, M. R., Griffin, R. S., Darney, D., Otte, K., & Ko, J. (2011). Urban African American youth exposed to community violence: A school-based anxiety preventive intervention efficacy study. *Journal of Prevention & Intervention in the Community*, 39(2), 149–166.
- Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.
- Cushing, L, Faust, J., August, L. M., Cendak, R., Wieland, W., & Alexeeff, G. (2015). Racial/Ethnic disparities in cumulative environmental health impacts in California: Evidence from a statewide environmental justice screening tool (CalEnviroScreen 1.1) *American Journal of Public Health*, 105(11), 2341–2348.
- *Decker, D. M., Dona, D. P., & Christenson, S. L. (2007). Behaviorally at-risk African American students: The importance of student–teacher relationships for student outcomes. *Journal of School Psychology*, 45(1), 83–109. <https://eric.ed.gov/?id=EJ748944>
- Deno, S. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review*, 15(3), 358–374. <https://eric.ed.gov/?id=EJ336942>
- Desimone, L., & Long, D. A. (2010). Teacher effects and the achievement gap: Do teacher and teaching quality influence the achievement gap between Black and White and high-and low-SES students in the early grades. *Teachers College Record*, 112(12), 3024–3073. <https://eric.ed.gov/?id=EJ913412>
- Earle, J., Maynard, R., Neild, R. C., Easton, J. Q., Ferrini-Mundy, J., Albro, E., et al. (2013). *Common guidelines for education research and development*. Washington, DC: Institute of Education Sciences. <https://eric.ed.gov/?id=ED544215>
- *Easton-Brooks, D., & Davis, A. (2009). Teacher qualification and the achievement gap in early primary grades. *Education Policy Analysis Archives*, 17(14/15), 1–16.

- *Eddy, C. M., & Easton-Brooks, D. (2011). Ethnic matching, school placement, and mathematics achievement of African American students from kindergarten through fifth grade. *Urban Education, 46*(6), 1280–1299. <https://eric.ed.gov/?id=EJ940675>
- *+Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review, 45*, 44–52.
- Every Student Succeeds Act of 2015. (2015). Pub. L. No. 114-95, 129 Stat. 1802, Title I, Section 8101(21)(A).
- *+Gershenson, S., Hart, C., Lindsay, C., & Papageorge, N. W. (2017). *The long-run impacts of same-race teachers* (IZA Discussion Paper No. 10630). Retrieved October 13, 2017, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2940620.
- *Gordon, D. M., Iwamoto, D. K., Ward, N., Potts, R., & Boyd, E. (2009). Mentoring urban Black middle-school male students: Implications for academic achievement. *Journal of Negro Education, 78*(3), 277–289. <https://eric.ed.gov/?id=EJ878480>
- Goza, F., & Ryabov, I. (2009). Adolescents' educational outcomes: Racial and ethnic variations in peer network importance. *Journal of Youth and Adolescence, 38*(9), 1264–1279. <https://eric.ed.gov/?id=EJ850943>
- *Grana, D. P. (2010). *Factors that influence eighth grade African American and White students' mathematics scores on the Missouri Assessment Program*. (Doctoral dissertation, University of Missouri, Saint Louis, 2010). Retrieved from ProQuest LLC. (UMI No. 3417226). <https://eric.ed.gov/?id=ED519260>
- *Graves, S. (2011). School and child level predictors of academic success for African American children in third grade: Implications for No Child Left Behind. *Race Ethnicity and Education, 14*(5), 675–697. <https://eric.ed.gov/?id=EJ947264>
- *Graves Jr., S. L., & Brown Wright, L. (2011). Parent involvement at school entry: A national examination of group differences and achievement. *School Psychology International, 32*(1), 35–48. <https://eric.ed.gov/?id=EJ916210>
- *Gutman, L. M., Sameroff, A. J., & Eccles, J. S. (2002). The academic achievement of African American students during early adolescence: An examination of multiple risk, promotive, and protective factors. *American Journal of Community Psychology, 30*(3), 367–399.
- *Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in context: School moderation of the impact of social identity threat on racial/ethnic achievement gaps. *Sociology of Education, 87*(2), 106–124. <https://eric.ed.gov/?id=EJ1021010>

- *Hemelt, S. W., & Marcotte, D. E. (2013). High school exit exams and dropout in an era of increased accountability. *Journal of Policy Analysis and Management*, 32(2), 323–349. doi:10.1002/pam.21688. <https://eric.ed.gov/?id=EJ1010255>
- Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher*, 33(8), 16–30.
- Horowitz, R., & Samuels, S. J. (Eds.). (2017). *The achievement gap in reading: Complex causes, persistent issues, possible solutions*. New York, NY: Routledge, Taylor & Francis Group. <https://eric.ed.gov/?id=ED574020>
- *Jackson, S. F. J. (2012). Self regulated and communal learning contexts as they relate to math achievement and math self efficacy among African American elementary level students. (Doctoral dissertation, Howard University, 2012). Retrieved from ProQuest LLC. (UMI No. 3591948). <https://eric.ed.gov/?id=ED559777>
- Jacobson Chernoff, J., Flanagan, K. D., McPhee, C., and Park, J. (2007). *Preschool: First findings from the preschool follow-up of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) (NCES 2008-025)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics. <https://eric.ed.gov/?id=ED498793>
- *Jeynes, W. H. (2005). The effects of parental involvement on the academic achievement of African American youth. *Journal of Negro Education*, 74(3), 260–274. <https://eric.ed.gov/?id=EJ750112>
- Kaminski, R. A., & Good, R. H. (2002). Letter naming fluency. In R. H. Good & R. A. Kaminski (Eds.), *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.
- *Kim, J. S. (2006). Effects of a voluntary summer reading intervention on reading achievement: Results from a randomized field trial. *Educational Evaluation and Policy Analysis*, 28(4), 335–355. <https://eric.ed.gov/?id=EJ759814>
- *Li, H. (2016). How is formative assessment related to students' reading achievement? Findings from PISA 2009. *Assessment in Education: Principles, Policy & Practice*, 23(4), 473–494. <https://eric.ed.gov/?id=EJ1119749>
- *Lleras, C., & Rangel, C. (2009). Ability grouping practices in elementary school and African American/Hispanic achievement. *American Journal of Education*, 115(2), 279–304. <https://eric.ed.gov/?id=EJ826476>
- *Martin, A. G. (2015). *The impact of flipped instruction on middle school mathematics achievement*. Unpublished doctoral dissertation, Texas A&M University, Commerce.

- McDaniel, M., Woods, T., Pratt, E., & Simms, M. C. (2017). *Identifying racial and ethnic disparities in human services: A conceptual framework and literature review*. OPRE Report #2017-69. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Retrieved on December 15, 2017 from: https://www.acf.hhs.gov/sites/default/files/opre/identifying_racial_and_ethnic_disparities_b508.pdf
- *Mezuk, B. (2009). Urban debate and high school educational outcomes for African American males: The case of the Chicago Debate League. *Journal of Negro Education*, 78(3), 290–304. <https://eric.ed.gov/?id=EJ878481>
- *Miranda, A., Webb, L., Brigman, G., & Peluso, P. (2007). Student Success Skills: A promising program to close the academic achievement gap for African American and Latino students. *Professional School Counseling*, 10(5), 490–497.
- Musu-Gillette, L., de Brey, C., McFarland, J., Hussar, W., Sonnenberg, W., & Wilkinson-Flicker, S. (2017). *Status and trends in the education of racial and ethnic groups 2017* (NCES 2017-051). Washington, DC: U.S. Department of Education, National Center for Education Statistics. <https://eric.ed.gov/?id=ED574873>
- *Nagle, B. T. (2013). *Out-of-school time program test score impact for Black children of single-parents*. (Doctoral dissertation, George Washington University, 2013). Retrieved from ProQuest LLC. (UMI No. 3555482). <https://eric.ed.gov/?id=ED552723>
- National Center for Education Statistics. (2016). *Table 219.46. Public high school 4-year adjusted cohort graduation rate (ACGR), by selected student characteristics and state: 2010-11 through 2014-15*. Retrieved October 13, 2017, from https://nces.ed.gov/programs/digest/d16/tables/dt16_219.46.asp.
- National Center for Education Statistics. (2017). *NAEP data explorer*. Retrieved October 13, 2017, from <https://nces.ed.gov/nationsreportcard/naepdata>.
- No Child Left Behind Act of 2001. (2002). Pub. L. No. 107-110, 115 Stat. 1425.
- Reardon, S. F., & Portilla, X. A. (2016). *Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry* (CEPA Working Paper No.15-02). Stanford, CA: Stanford University, Center for Education Policy Analysis. Retrieved October 10, 2017, from <http://cepa.stanford.edu/wp15-02>.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy, and Practice*, 5, 77–84.

- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, 34, 379–440.
- *Strunk, K. O., & McEachin, A. (2014). More than sanctions: Closing achievement gaps through California's use of intensive technical assistance. *Educational Evaluation and Policy Analysis*, 36(3), 281–306. <https://eric.ed.gov/?id=EJ1042020>
- U.S. Department of Education. (2016). *Non-regulatory guidance: Using evidence to strengthen education investments*. Washington, DC: Author. Retrieved from <https://www2.ed.gov/policy/elsec/leg/essa/guidanceusesinvestment.pdf>
- *Wang, A. H. (2010). Optimizing early mathematics experiences for children from low-income families: A study on opportunity to learn mathematics. *Early Childhood Education Journal*, 37(4), 295–302. <https://eric.ed.gov/?id=EJ867587>
- *Webb, L. D., Brigman, G. A., & Campbell, C. (2005). Linking school counselors and student success: A replication of the Student Success Skills approach targeting the academic and social competence of students. *Professional School Counseling*, 8(5), 407–413. <https://eric.ed.gov/?id=EJ714284>
- *Wenglinsky, H. (2004a). Closing the racial achievement gap: The role of reforming instructional practices. *Education Policy Analysis Archives*, 12(64), 1–22. <https://eric.ed.gov/?id=EJ853528>
- *Wenglinsky, H. (2004b). The link between instructional practice and the racial gap in middle schools. *RMLE Online: Research in Middle Level Education*, 28(1), 1–13. <https://eric.ed.gov/?id=EJ807416>
- What Works Clearinghouse. (2014). *Procedures and standards handbook (Version 3.0)*. Retrieved October 13, 2017, from <https://eric.ed.gov/?id=ED544775>
- *White, L. J., Hixson, N., Hammer, P. C., Smith, D. L., & D'Brot, J. (2010). *Examining the effectiveness of closing the achievement gap professional development demonstration schools*. Charleston, WV: West Virginia Department of Research, Division of Curriculum and Instruction. <https://eric.ed.gov/?id=ED514364>
- Wisconsin Department of Public Instruction. (2014). *Promoting excellence for all: A report from the state superintendent's task force on Wisconsin's achievement gap*. Madison, WI: Author. Retrieved October 13, 2017, from <https://dpi.wi.gov/sites/default/files/imce/excforall/exc4all-report.pdf>.
- Wixom, M. A. (2015). *Closing the achievement gap: Four states' efforts*. Denver, CO: Education Commission for the States. <https://eric.ed.gov/?id=ED561914>

*Woolley, M. E., Strutchens, M. E., Gilbert, M. C., & Martin, W. G. (2010). Mathematics success of Black middle school students: Direct and indirect effects of teacher expectations and reform practices. *Negro Educational Review*, 61(1-4), 41-59. <https://eric.ed.gov/?id=EJ908042>



REL Midwest at American Institutes for Research
10 South Riverside Plaza, Suite 600
Chicago, IL 60606
312.288.7600 | RELMidwest@air.org