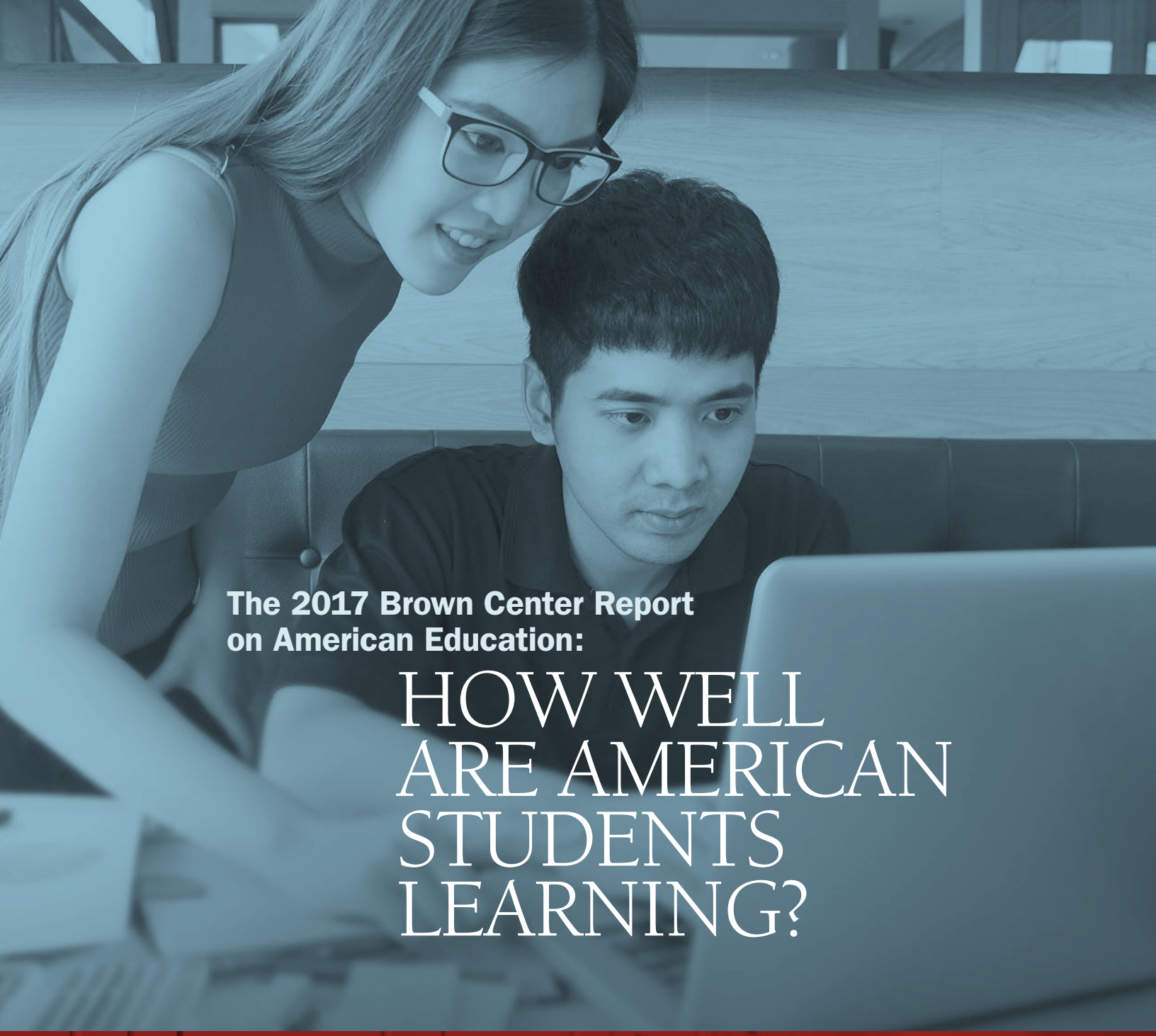


The 2017 Brown Center Report on American Education:

HOW WELL ARE AMERICAN STUDENTS LEARNING?

With sections on the latest international test scores, foreign exchange students, and school suspensions

BROWN CENTER on Education Policy at BROOKINGS

ABOUT BROOKINGS

The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars. Brookings recognizes that the value it provides is in its absolute commitment to quality, independence, and impact. Activities supported by its donors reflect this commitment.

ABOUT THE BROWN CENTER ON EDUCATION POLICY

Raising the quality of education in the United States for more people is imperative for society's well-being. With that goal in mind, the purpose of the Brown Center on Education Policy at Brookings is to examine the problems of the American education system and to help delineate practical solutions. For more information, see our website, www.brookings.edu/brown.

This report was made possible by the generous financial support of The Brown Foundation, Inc., Houston. Appreciation is also extended to AFS-USA, Inc. for assistance in conducting the report's survey of international students.

The 2017 Brown Center Report on American Education:

HOW WELL ARE AMERICAN STUDENTS LEARNING?

With sections on the latest international test scores, foreign exchange students, and school suspensions

March 2017
Volume 3, Number 6

by:
TOM LOVELESS
Nonresident Senior Fellow, The Brown Center on Education Policy,
The Brookings Institution

TABLE OF CONTENTS

Introduction

This Brown Center Report (BCR) on American Education is the sixth and final edition in the third volume and the 16th issue overall. The series began in 2000. As in the past, the report comprises three studies. Also in keeping with tradition, the first section features recent results from state, national, or international assessments; the second section investigates a thematic topic in education, either by collecting new data or by analyzing existing empirical evidence in a novel way; and the third section looks at one or more education policies.

In Part I, this year's focus is on the latest results from two international tests, the Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS). Both tests were administered in 2015, and the U.S. participated in both. TIMSS tests fourth and eighth grade students in math and science. PISA tests 15-year-olds in reading literacy, mathematics literacy, and science literacy. TIMSS began in 1995. From 1995–2015, the U.S. made statistically significant gains on TIMSS fourth grade math, eighth grade math, and eighth grade science assessments. The four-point scale score gain in fourth grade science is not statistically significant. PISA began in 2000. Since PISA's inception, U.S. scores have been flat on all three subjects; however, the 2015 math score of 470 marks a significant decline from 481 in 2012 and 487 in 2009.

Part II revisits one of the most popular studies in BCR history, a 2001 survey of foreign exchange students. The same survey was conducted in 2016. The idea is simple, asking kids from abroad who have attended U.S. high schools what they think about U.S. education and their American peers. Comparing the results, 15 years apart, suggests that not much has changed. International students still think U.S. schools are much less

challenging than schools in their home countries and that American teens are more focused on success at sports compared to their peers back home.

Part III examines race and school discipline. Exclusionary punishments, those that remove students from schools, have come under fire in recent years. California officials have been pushing schools to reduce out-of-school suspensions, especially because of the racial disparities associated with that form of discipline. The policy has succeeded in reducing suspensions in the state—they are down dramatically—but racial disparities persist. Black students continue to be suspended at three to four times their proportion of student enrollment. The study examines three years (2013–2015) of California school-level data (a sample of 7,180 schools) to identify characteristics of schools that are correlated with high- and low-suspension-rates for African-American kids.

Schools with large populations, schools exclusively serving middle school grades (e.g., grades six to eight), and schools serving a high proportion of poor or black students are all associated with elevated suspension rates for African-Americans. Disciplinary reformers have promoted restorative programs as alternatives to exclusionary punishment, but the approaches are controversial and the empirical evidence of their impact is limited. The current study cannot draw causal conclusions, but altering the structural characteristics of schools associated with higher suspension rates should be considered in future reform efforts.

Part I INTERNATIONAL ASSESSMENTS

THE UNITED STATES PARTICIPATES IN TWO INTERNATIONAL assessments: the Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS). The latest scores for both tests, conducted in 2015, were released in December 2016. PISA is given to 15-year-old students every three years and TIMSS to fourth and eighth graders every four years, meaning that the two tests coincide every 12 years. 2015 was such a year. In addition to allowing for a comparison of U.S. students to students in other countries, scores from PISA and TIMSS join the National Assessment of Educational Progress (NAEP) as providing the only valid estimates of U.S. national academic performance. The three tests use similar sampling designs, allowing for the results from tested samples to be generalized to the entire nation.

PISA was first administered in 2000. Reading literacy was the major subject assessed that year, followed by mathematics literacy in 2003 and science literacy in 2006. Table 1-1 displays U.S. scores for all PISA assessments between 2000 and 2015. Like TIMSS, PISA is scaled with an international mean of 500 and standard deviation of 100. U.S. scores have been flat over the 15 years of PISA. Reading literacy scores have hovered near the international mean, ranging between 495 (in 2003) and 504 (in 2000). Mathematics literacy scores have come in below the mean, with U.S. scores ranging between 470 and 487. The most recent mathematics score, 470, is the lowest for the U.S. in the test's history. Science literacy scores have also fluctuated within a tight range, 489–502. The lack of an asterisk next to the 2015 scores means that none of the three PISA subjects registered a statistically significant change between the year when they were introduced and 2015. The 2015 math score is statistically significantly lower than the scores of 2009 and 2012, however.

TIMSS fourth grade scores are shown in Table 1-2. Compared to the PISA scores, the U.S. performs better on TIMSS, both in terms of absolute levels and in gains over time. Math scores have stayed solidly above the international mean of 500 for the entire

20-year period of 1995–2015, and the latest score of 539 represents a statistically significant gain from the score of 518 in 1995. Science scores have held in a narrow channel, from 536 to 546. The gain of four points from 1995–2015 is not statistically significant.

On the eighth grade TIMSS, the U.S. notched statistically significant gains in both math and science over the course of TIMSS history (see Table 1-3). Math scores rose from 492 to 518. Science scores rose from 513 to 530. In contrast to PISA, the U.S. performs significantly above the international mean in math and science, with scores on the upswing. The PISA-TIMSS difference is especially surprising when one considers that the 15-year-olds taking the PISA exam and eighth graders taking TIMSS are not far apart in their school careers. About 70% of students in the PISA sample are in the fall semester of their sophomore year (10th grade) of high school. The TIMSS eighth grade sample is tested in the spring. For at least seven out of 10 examinees in the PISA sample, then, students have had the entire ninth grade, a couple of months of eighth and 10th grades, and two intervening summers since they were eligible for the TIMSS sample. That is not a lot of schooling to differentiate the two groups.

U.S. scores on PISA, 15-year-olds (2000–2015) — TABLE 1-1

	2000	2003	2006	2009	2012	2015
Reading	504	495	—	500	498	497
Math		483	474	487	481	470
Science			489	502	497	496

Source: PISA 2015 Results (Volume I) Excellence and Equity in Education, Table I.4a (Reading); Table I.2.4a; Table I.5.4a (Math).

U.S. scores on TIMSS, 4th Grade (1995–2015) — TABLE 1-2

	1995	2003	2007	2011	2015
Math	518	518	529	541	539*
Science	542	536	539	544	546

Source: Highlights From TIMSS and TIMSS Advanced 2015, NCES, Figure 2a. (Math); Figure 6a. (Science)
Note: A "*" indicates a statistically significant change between the 1995 score and the 2015 score ($p<0.05$).

U.S. scores on TIMSS, 8th Grade (1995–2015) — TABLE 1-3

	1995	1999	2003	2007	2011	2015
Math	492	502	504	508	509	518*
Science	513	515	527	520	525	530*

Source: Highlights From TIMSS and TIMSS Advanced 2015, NCES, Figure 2b. (Math); Figure 6b. (Science).
Note: A "*" indicates a statistically significant change between the 1995 score and the 2015 score ($p<0.05$).

Why the handwringing? International comparisons.

According to PISA, U.S. school performance has been flat for 15 years. TIMSS paints a rosier picture, with significant gains in fourth and eighth grade mathematics and eighth grade science. A flat to positive trend does not seem to justify handwringing over U.S. performance. Yet handwringing about how the U.S. does on international tests contends with baseball as a national pastime. Former Secretary of Education Arne Duncan called the 2009 PISA results "an absolute wakeup call,"[1] a comment that seemed strangely ahistorical at the time considering the U.S. scored much worse—11th out of 12 countries—in the First International Mathematics Study (FIMS), administered more than four decades earlier in 1964.[2]

The despair arises from how the U.S. compares to economically developed countries in Europe and Asia. Despite gains on TIMSS, the U.S. still scores far below the top performers. Singapore provides a good comparison because it scored the

Rankings are simple to understand. They also can mislead.

TIMSS scores, Singapore and U.S. (2015) TABLE 1-4

2015 TIMSS	4th Grade		8th Grade	
	Math	Science	Math	Science
Singapore	618	590	621	596
U.S.	539	546	518	530

Source: Highlights From TIMSS and TIMSS Advanced 2015, NCES, Figure 1a. (4th grade math); Figure 1b. (8th grade math); Figure 5a. (4th grade science); Figure 5b (8th grade science).

U.S. PISA scores (2015), in relation to other participants TABLE 1-5
(Number of systems)

U.S. score is...	Reading (23T)	Math (39T)	Science (25)
Lower than	14	36	18
Statistically indistinguishable from	13	5	12
Higher than	42	28	39

Note: The number in parentheses represents the official ranking of the U.S. on the assessment; a "T" indicates that the U.S. tied for that ranking.

U.S. TIMSS Scores (2015), in relation to other participants TABLE 1-6
(Number of systems)

U.S. score is...	4th Grade		8th Grade	
	Math (13T)	Science (9)	Math (9T)	Science (10)
Lower than	10	7	8	7
Statistically indistinguishable from	9	7	11	9
Higher than	34	38	24	26

Note: The number in parentheses represents the official ranking of the U.S. on the assessment; a "T" indicates that the U.S. tied for that ranking.

highest on the 2015 TIMSS math and science assessments at both grade levels (see Table 1-4). The TIMSS scale theoretically runs from 0–1,000, but as an empirical matter, scores range from the 300s to the 600s. As shown in Table 1-4, the U.S. lags Singapore by at least 44 points (fourth grade science)—and by 103 points in eighth grade math! That difference is a full standard deviation. The good news is that the U.S.-Singapore eighth grade math gap has narrowed since 1995 (when it was 117 points); the bad news is that it will take, at this pace, more than 140 years to close it completely.

Researchers shy away from using rankings in serious statistical analyses of test scores, but they are frequently used in political advocacy, most visibly in media headlines or sound bites. Rankings are simple to understand and conjure up the image of team standings in a sports league. They also can mislead. National scores on TIMSS and PISA are estimates, bounded by confidence intervals that reflect sampling error. Sampling error is not really "error" in the common sense of the word, but statistical noise introduced by inferring national scores from a random sample of test takers. Because every nation's score is estimated in this way, it cannot be said with confidence that the rankings of participants with overlapping confidence intervals actually differ; they are considered statistically indistinguishable.

When new scores are released, the National Center of Educational Statistics (NCES) does its best to provide an accurate summary of U.S. rankings on PISA and TIMSS. It does so by describing the U.S. relative performance while taking statistical significance into consideration. Table 1-5 presents the 2015 PISA data in a similar fashion. The PISA scores are still disappointing, but not as dramatically as they initially seem. The reading scores, in particular, illustrate the nebulousness of rankings. The U.S. score in reading is tied for 23rd place, but its true ranking is more complicated than that. When statistical significance is taken into account, 14 systems scored higher than the U.S, 13 scored about the same, and 42 scored lower.

The U.S. also looks better on TIMSS (see Table 1-6) when scores are considered in this context. On the fourth grade

TIMSS test in mathematics, the U.S. score is reported as tied for 13th place. More precisely, it scores below 10 systems, is statistically indistinguishable from the scores of nine systems, and is higher than the scores of 34 systems.[3] In eighth grade math, the contrast with PISA's math scores is provocative. Only eight systems outscore the U.S. on TIMSS, compared to 36 countries outscoring the U.S. on PISA math. Five countries that scored significantly lower than the U.S. on TIMSS—Australia, Sweden, Italy, Malta, and New Zealand—scored significantly higher than the U.S. on PISA.

National test score correlations: TIMSS and PISA

Previous Brown Center Reports have discussed key differences of TIMSS and PISA. TIMSS is grade-based, and PISA is age-based.[4] TIMSS tests fourth and eighth graders, while PISA tests 15-year-olds. TIMSS is curriculum-based, meaning that it measures how well students have learned reading, mathematics, and science as presented in the school curriculum. PISA is a test of how well students can apply what they have learned to solve real world problems (hence "literacy" appended to the common labels for school subjects) and reflects what PISA's expert committees believe students should know or need to know.[5]

Despite these differences, TIMSS and PISA test scores are highly correlated. Table 1-7 displays the correlation coefficients for 2015 TIMSS and PISA scores. All three PISA tests are strongly correlated. The surprise is the magnitude of the correlation of PISA's reading test with both math (0.91) and science (0.96). The two TIMSS tests are also highly correlated (0.92). And, as expected, PISA's math scores are highly correlated with TIMSS math scores (0.93)—and PISA science scores with TIMSS science scores (0.94).

Correlation coefficients, TIMSS (2015) and PISA (2015) — TABLE 1-7

	PISA-Reading	PISA-Science	TIMSS-Math	TIMSS-Science
PISA-Math	0.91	0.97	0.93	0.94
PISA-Reading		0.96	0.78	0.90
PISA-Science			0.87	0.94
TIMSS-Math				0.92

Note: N = 27 countries participating in both TIMSS 2015 (8th grade) and PISA 2015.

Researchers have drawn different implications from these correlations. Economists Eric Hanushek and Ludger Woessmann concluded that the two tests measure "a common dimension of skills," and that the scores can be aggregated to form a single national-level indicator of cognitive ability predicting economic growth.[6] Psychologist Heiner Rindermann referred to that common dimension as a "g-factor," standing for general intelligence. The term touches upon a longstanding debate in psychology. Simply put, the argument is about the extent to which human intelligence is general (smart people are smart about most things) or specific (smart people in math are not necessarily smart in interpreting poetry).[7]

Eckhard Klieme, an educational researcher with intimate knowledge of TIMSS and PISA, examined 2015 data for both TIMSS and PISA math assessments and analyzed the tests' correlations. Klieme acknowledges that the tests' cross-sectional scores are highly correlated but he also explores differences. He shows, for example, that the small differences between scores from the two tests can be explained by content coverage, the topics that math teachers reported being taught. Countries in which teachers reported teaching more of the TIMSS content scored higher on the TIMSS

What is a Correlation Coefficient?

A Pearson correlation coefficient measures the strength of a linear relationship between two variables. The coefficient is always between –1.00 and +1.00. The closer a coefficient is to +/–1.00 the stronger a relationship is between two variables. 1.00 signifies a perfect positive relationship while –1.00 signifies a perfect negative relationship.

On the 2015 PISA, the U.S. continued to register mediocre scores; on TIMSS it does better.

test than would be predicted from their PISA score. He also found that gain scores from the two tests were not as strongly correlated, with a 0.61 correlation of PISA and TIMSS gains from 2003 to 2015. That is strong but substantially weaker than the cross-sectional correlations for 2015.[8]

In the current study, a total of 22 systems participated in 2011 and 2015 TIMSS (eighth grade math) and 2012 and 2015 PISA. The correlation coefficient for their TIMSS and PISA math gains is 0.52. That, too, is much weaker than the cross-sectional correlations reported in Table 1-7.

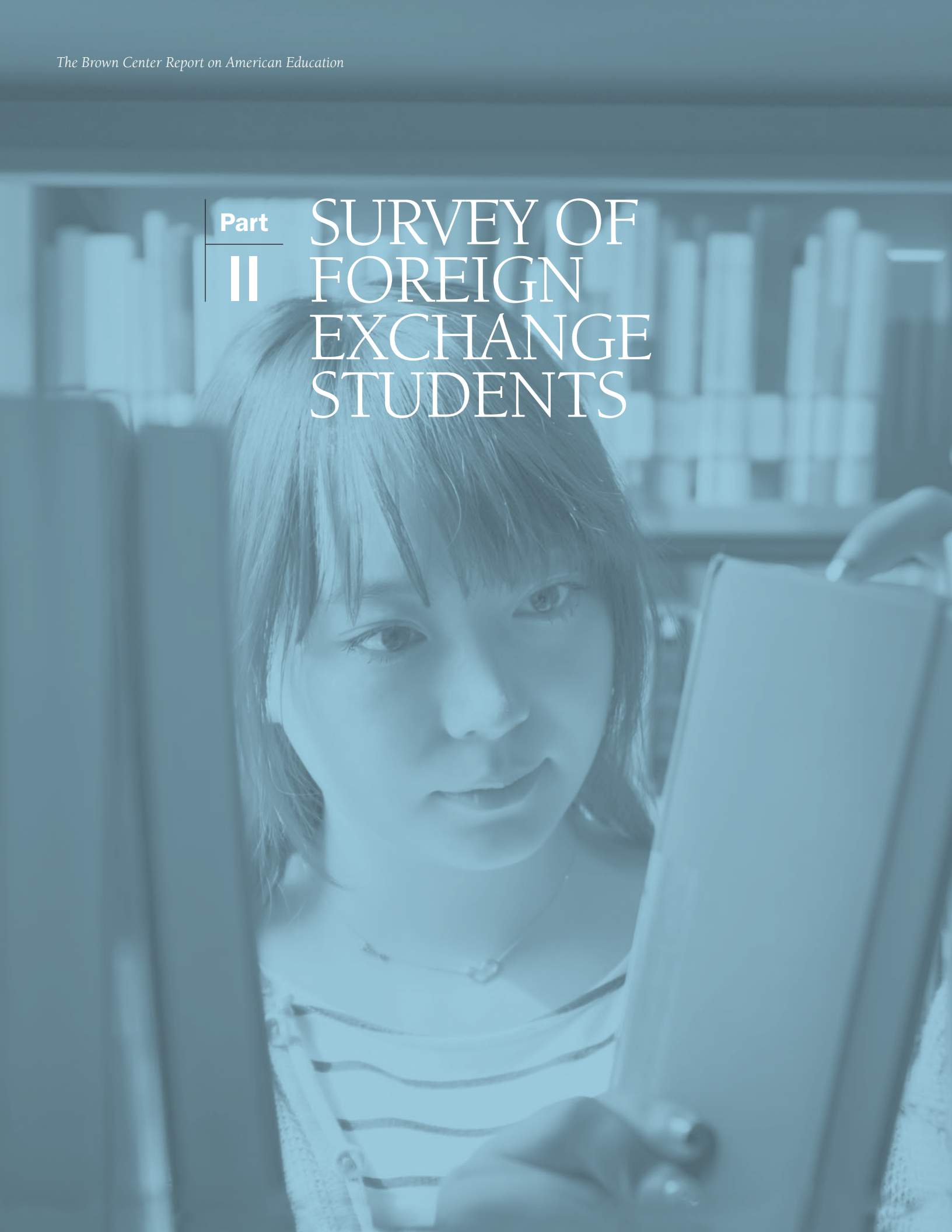
Conclusion

On the 2015 PISA, the U.S. continued to register mediocre scores, as it has done since PISA began in 2000. The mathematics literacy score of 470 represented a statistically significant decline of 11 scale score points from the 481 scored in 2012, but U.S. performance in all three subjects—math, reading, and science—was not statistically significantly different from how the nation performed when each subject was first administered. TIMSS scores were more encouraging for the U.S., especially at the eighth grade level, where statistically significant gains have been made in both math and science since 1995. Significant gains on TIMSS have also been made in fourth grade math since 1995.

PISA and TIMSS scores are highly correlated. Cross-sectional test scores are often highly correlated when aggregated to the state or national level. It is important to note what these high correlations do not mean. They do not mean that the tests assess the same knowledge or skills; otherwise, countries are wasting a lot of time giving three PISA tests when the PISA reading literacy test is a good tool for measuring achievement in science (r = 0.96) and math (r = 0.91). High-achieving 10th graders in the U.S. take an advanced algebra course in mathematics. Imagine administering a reading test to see how well they learned algebra!

Casual observers of international tests should pay close attention to the trends on both tests. As shown above, PISA and TIMSS trend data are not as strongly correlated as the cross-sectional scores. The U.S. is showing steady progress on TIMSS but scores are flat on PISA—even declining in mathematics on the last two rounds.

Comparing the U.S. with other countries must be done with caution. Finland scored among the top countries on PISA in the early 2000s and became a famous destination for American "edutourists" eager to visit Finland's schools. Since 2006, Finland's PISA scores have declined dramatically. On TIMSS, fourth grade math scores for Finland (535) and the U.S. (539) are statistically indistinguishable. Speaking in Washington, D.C. in 2010, Organisation for Economic Co-operation and Development (OECD) Secretary-General Angel Guerria called New Zealand a "top flier," and one of the "strongest overall performers."[9] And yet, since 1995, New Zealand has consistently scored either at comparable levels or below the U.S. on TIMSS—in both math and science and at both the fourth and eighth grade levels. More importantly, New Zealand's TIMSS scores have been falling during the last several rounds of TIMSS, while the U.S. scores have been climbing. To get the most value from U.S. participation in PISA and TIMSS, policymakers—and the public—should pay close attention to the trends on both tests.



Part II SURVEY OF FOREIGN EXCHANGE STUDENTS

IN THE FALL OF 1957, JAMES COLEMAN BEGAN SURVEYING students in 10 high schools in northern Illinois. The sample was not randomly drawn. Schools were selected to reflect varying sizes (150 to 1,850) and communities (small towns, suburbs, and cities), serving students whose parents worked at varying occupations (farming, industrial, and professional). One of the schools was an all-boys Catholic school; the rest were public schools.

Coleman was careful to acknowledge that the sample was not designed to be representative of high schools in general, or even of high schools in northern Illinois, and that the results could not be generalized to a larger population. Despite these limitations, when the findings appeared in "The Adolescent Society," published in 1961, the book was instantly recognized as a classic in the study of education.

Coleman's unique insight was that modernity had given birth to a social unit unknown to previous generations, an adolescent subculture with its own values, norms, language, and status system. Although the 20th century had dawned with most Americans leaving school for work after completing eighth grade (only 11% of high-school-aged teens in 1900 actually attended high school), by the 1950s high school attendance had become nearly universal. Unlike families of agrarian societies, the modern family no longer served as society's primary economic unit. The modern high school walled off adolescents from adult society and, by extending the length of time for youths' education, delayed entry into adulthood. Consequently, Coleman argued, this age-segregated world "comes to constitute a small society, one that has most of its important interactions *within* itself, and maintains only a few threads of connection with the outside adult society."[10]

Coleman found high schools remarkably anti-academic. Peer status underpinned a caste-like system with athletes at the top and brilliant students further down the social hierarchy. Coleman noted the irony of academic learning receiving so few status rewards among teens. The main reason children go to school, as reflected in compulsory education laws, is to learn, not to play sports.

The current study

In the six decades since Coleman's study, education researchers have frequently surveyed teens and asked them about their school experiences. In 2001, it dawned on researchers in the Brown Center that one particular group of teens, foreign exchange

students, had never been singled out for special consideration. We believed this group of teenagers could offer a fresh perspective on U.S. high schools. They are also teenagers; most of them attend modern high schools populated by children of a similar age, and with the worldwide diffusion of adolescent culture, they surely share some of the same views as American teens. Perhaps they could also shed light on what is peculiarly American about American high schools.

In the spring of 2001, we conducted a survey of students from abroad who were nearing the end of their academic year in U.S. schools. The results appeared in the 2001 Brown Center Report. In the summer and fall of 2001, we replicated the survey with U.S. students who had gone abroad to attend high school, asking them the same set of questions. Those results appeared in the 2002 Brown Center Report.

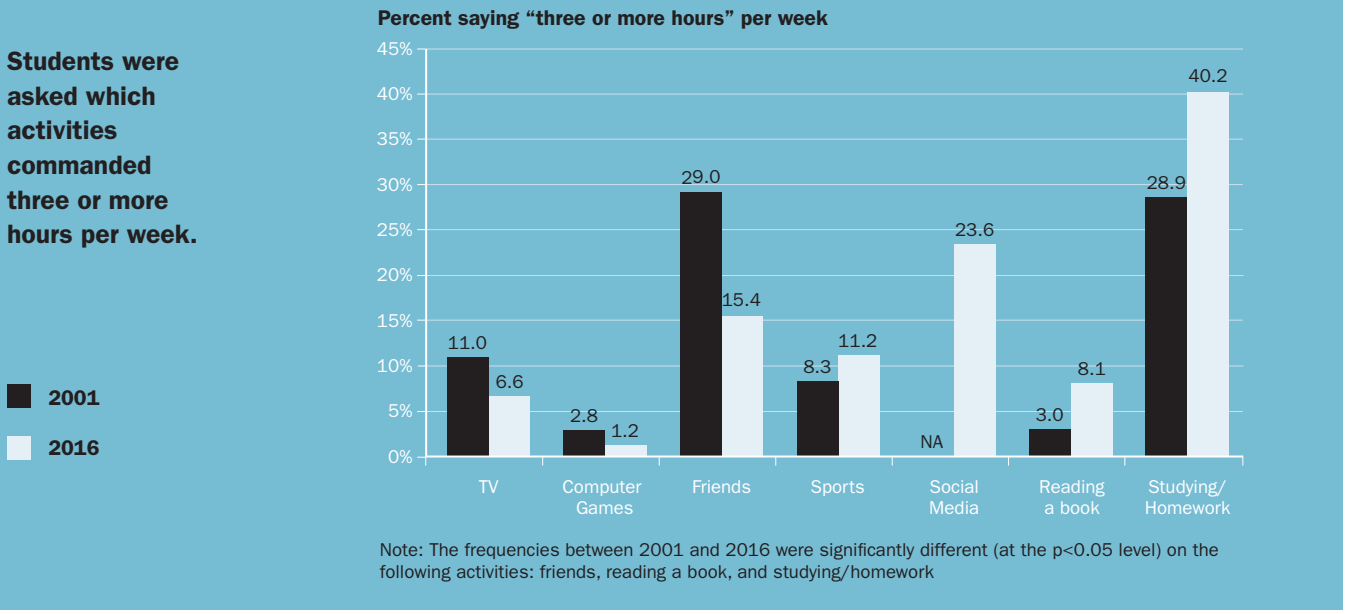
It is time to replicate at least a portion of that study. In the spring of 2016, a survey of foreign exchange students was conducted. The same set of questions asked in 2001, with a few modifications, was asked once again. Has anything changed over the past 15 years? A random sample of 600 foreign exchange students was drawn from a list of all international students attending U.S. high schools under the auspices of AFS International. The survey was conducted by mail. Responses were received from 259 students (28 surveys were returned as undeliverable) for a response rate of 45.3%.

In the analysis below, the discussion will focus on the impressions of foreign exchange students regarding U.S. education and how those impressions have changed since the surveys that were conducted in 2001.

Time spent on out-of-school activities

FIGURE 2-1

Students were asked which activities commanded three or more hours per week.



Note: The frequencies between 2001 and 2016 were significantly different (at the p<0.05 level) on the following activities: friends, reading a book, and studying/homework

Homework or studying is the dominant activity outside of school.

What do students do outside of school?

We asked the international students how they spend time in their home country when they are not in school. Figure 2-1 shows the out-of-school activities that students spend more than three hours per week doing. Homework or studying is the dominant activity. Four out of 10 students (40.2%) report spending at least three hours studying or doing homework each week, significantly more than 28.9% with the same response in 2001. Social media, which was not included on the 2001 survey, also gets a lot of time, with 23.6% of respondents naming it as a popular activity. Note that the percentage of students saying they spend three or more hours with friends dropped sharply from 2001 to 2016 (29.0% to 15.4%). A significant portion of teens' social lives is now online.

More international students devoted three hours a week reading books for pleasure in 2016 (8.1%) than in 2001 (3.0%). That trend is the opposite of an apparent decline in reading for pleasure among U.S. teens.[11] On the 2015 NAEP, 37% of U.S. 12th graders said they never read for pleasure, up from 30% in 2005.[12] That raises an important point: the percentage of foreign exchange students saying they spent "no time" reading for pleasure (not shown in the figure) also increased—from 26.4% to 31.3%. The increase is not statistically significant, but it does show that, as in the U.S., a large proportion of teens abroad do not read for pleasure at all.

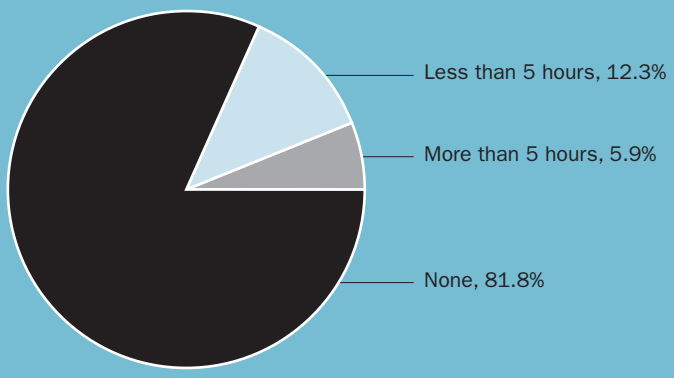
Part-time jobs

Whether students should work during high school has long been debated. Light (1995) analyzed data from the National

Time spent on work

FIGURE 2-2

Students were asked how many hours per week they spent working.



Longitudinal Survey of Youth (NLSY) and discovered a wage benefit from high school employment of about 10% that lasts approximately five years after high school graduation but then dissipates. Does part-time employment affect student performance at school? Marsh and Kleitman's (2005) analysis of the National Education Longitudinal Study of 1988 (NELSS88) found negative effects of working during high school on achievement, coursework selection, educational and occupational aspirations, and college attendance. Kalenkoski and Pabilonia (2011), analyzing data from the American Time Use Surveys, found that employment reduces the amount of time students devote to homework, with one hour spent on work associated with 11 fewer minutes spent on homework.

Compared to teens abroad, American teens are unique in working part time during high school. Studies in the 1990s indicated that about two-thirds to three-quarters of U.S. high school students were employed at some point during their high school careers, much more than frequencies found in other countries. That contrast was apparent in the Brown Center Report's 2001 surveys. Only 20% of students from abroad reported holding down a part time job while attending high school in their home countries. Compare that to U.S students who went abroad as exchange students: Slightly more than half (55%) said they worked part time.

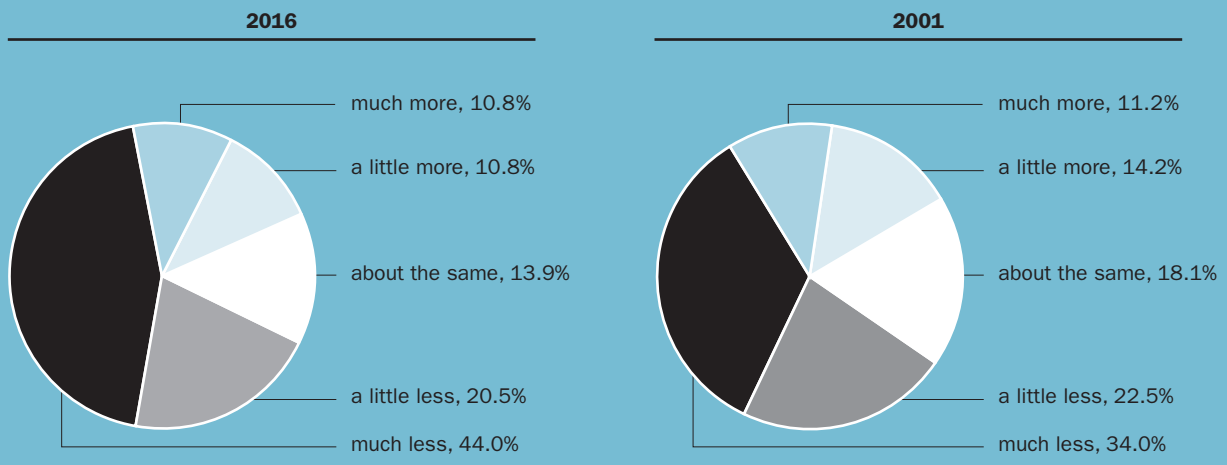
International students surveyed in 2016 indicate employment rates comparable to their counterparts 15 years ago (see Figure 2-2). Almost eight out of 10 (81.8%) do not work and only 5.9% spend more than five hours working at a job. In the 2001 survey of American students who studied abroad, more than a third (35%) said they spent more than five hours per week working.

U.S. exchange students are the appropriate comparison group, but we did not survey them in 2016. The 2001 statistics

Foreign exchange students' perceptions of time spent on schoolwork

FIGURE 2-3

Students were asked: Compared to students in your home country, do you think U.S. students spend more, less, or about the same amount of time on schoolwork?



International students think American students devote less time to schoolwork.

may need updating. Recent data from the Bureau of Labor Statistics suggest that only about half as many American teens are now working compared to 2000, a downward trend that accelerated during the Great Recession. Another caveat is that regulations that govern youth employment and place restrictions on the number of hours young people are allowed to work may have changed since 2001.[13]

Time spent on schoolwork

The survey asked students the following: Compared to students in your home country, do you think U.S. students spend more, less, or about the same amount of time on schoolwork? Responses from both 2001 and 2016 are displayed in Figure 2-3. International students think American students devote less time to schoolwork. In 2001, 34.0% said much less, a figure that grew to 44.0% in 2016. When the 20.5% who answered "a little less" are also considered, it means that nearly two-thirds of foreign exchange students (64.5%) believe U.S. high school students spend less time on schoolwork than their peers do back home.

Are U.S. classes easier or harder?

In the 2001 survey, foreign exchange students reported that high school classes in the U.S. seemed easier than classes in their home countries. When asked to rate the relative difficulty of U.S. classes, 56% replied "a lot easier" and 29% said "a little easier." Only 6% said "a little harder" and 5% said "much harder." The 2001 American students who had spent time abroad as foreign exchange students agreed, although less emphatically: 29% answered "much easier" and 27% said "a little easier" when asked how their classes at home compared to those abroad. Of the American students, 13% called U.S. classes "a little harder" and 17% "much harder." On average, then, American students also judged U.S. classes as easier.

Foreign exchange students' perceptions of class difficulty

FIGURE 2-4

Students were asked how U.S. classes compared to their courses at home.

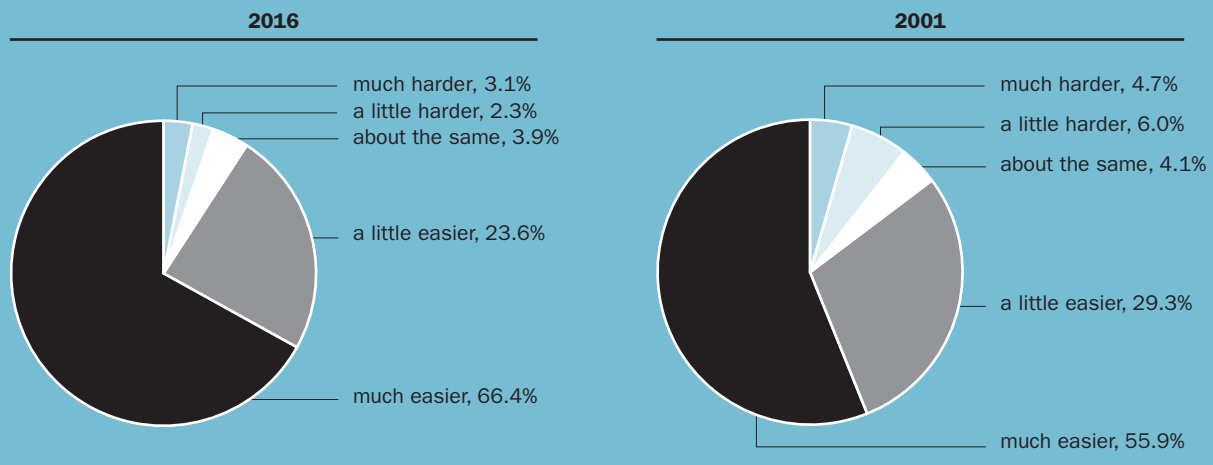


Figure 2-4 compares the responses of international students from 2001 and 2016. Students from abroad are even more likely today to describe U.S. classes as easier than they were in 2001. The combined "much easier" and "a little easier" responses grew from 85.2% in 2001 to 90.0% in 2016. The change in the "much easier" rating, increasing from 55.9% to 66.4%, is statistically significant. Considering the rhetoric of U.S. curriculum reform over the past 15 years—the calls for higher standards, more rigorous coursework, deeper learning, and stronger preparation for college—these results are surprising.

But not completely so. The U.S. has not acted alone in pursuing school reform. Making education more rigorous has been a policy objective of countries all over the world; therefore, even if high school coursework has become more challenging in the U.S., it may still appear less daunting compared to the academic demands of high schools in other countries.

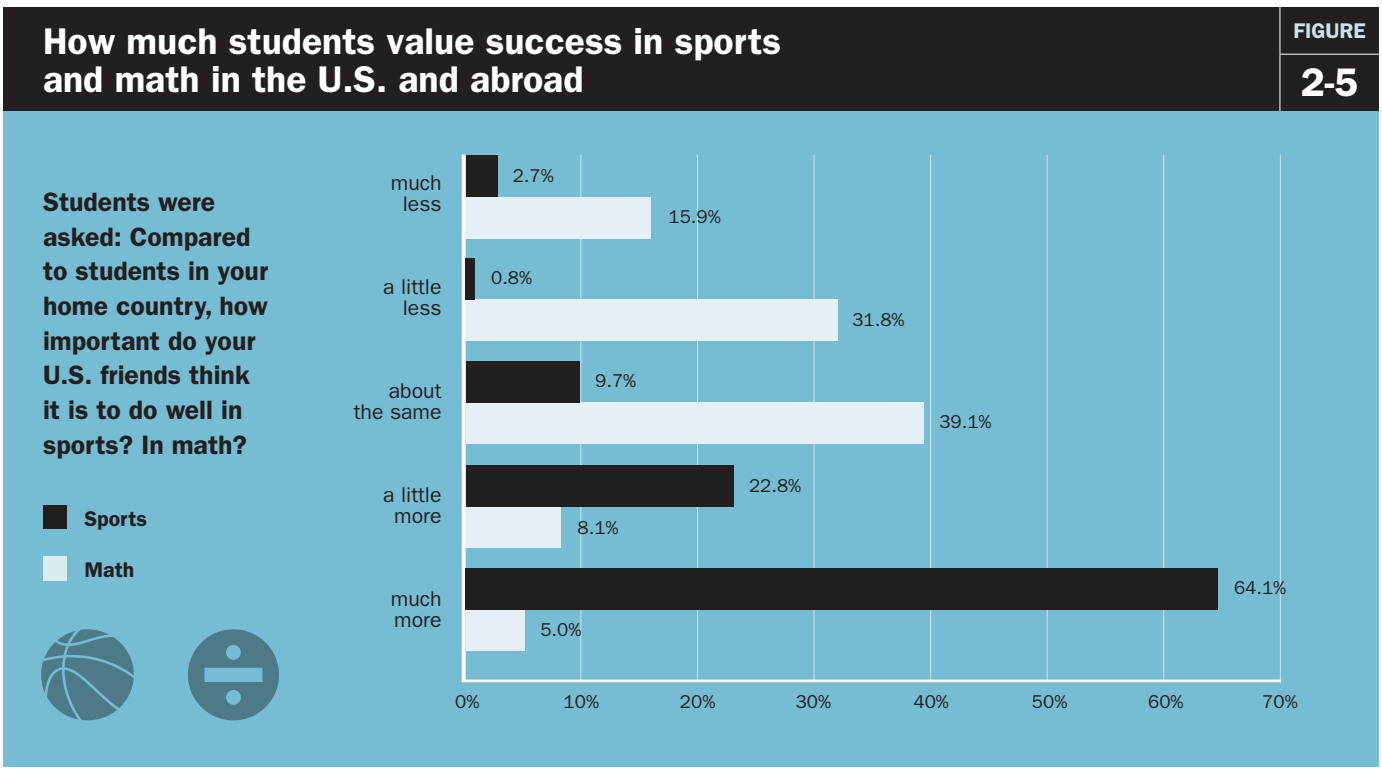
Valuing success at math and sports

One of the most intriguing findings from the 2001 surveys involved the importance peers attribute to success at math and sports. Students from abroad reported that American teens were much more likely to value success at sports than at math. The questions were asked again in 2016 and generated similar results.

Students were first asked: "Compared to students in your home country, how important do your U.S. friends think it is to do well in math?" Then students were asked to make the same comparison in regards to doing well in sports.

How much students value success in sports and math in the U.S. and abroad

FIGURE 2-5



Athletic success is just as important to U.S. teens today as it was in the mid-20th century.

Figure 2-5 presents the 2016 data. About one out of six (15.9%) respondents describe success at mathematics as "much less" important to friends in the U.S. compared to their friends back home. Almost one-third of students (31.8%) from abroad believe it is "a little less" important. The corresponding figures for sports are trivial. Only 2.7% describe sports as "much less" important in the U.S. and 0.8% as "a little less" important. The other end of the continuum, identifying what kind of accomplishments are important, produces a dramatic contrast. Nearly two-thirds of foreign exchange students (64.1%) see American teens as valuing success at sports "much more" than teens in the international students' home countries. For math, the corresponding figure is 5.0%. Another 22.8% rate sports as "a little more" valued in the U.S. The two categories indicating greater esteem accorded by U.S. teens to success in sports overwhelm the other response categories. The importance of athletic success in allocating status appears just as alive and powerful today in U.S. high schools as it was in the mid-20th century when Coleman conducted his research.

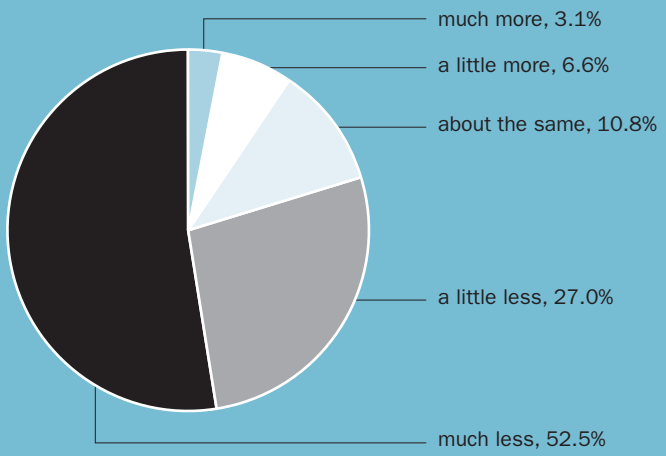
Learning a second language

The 2016 survey included a question not asked in 2001. Using the same response categories as the math and sports questions, the question asked students to evaluate the relative importance of learning a foreign language. The question was: "Compared to students in your home country, how important do your U.S. friends think it is to be able to speak a second language?" More than half (52.5%) of respondents said it was "much less" important in the U.S., with another 27.0% describing acquisition of a second language as "a little less" important.

How much students value learning a second language in the U.S. and abroad

FIGURE 2-6

Students were asked: Compared to students in your home country, how important do your U.S. friends think it is to be able to speak a second language?



This result is unsurprising in the sense that the survey population, after all, consists of foreign exchange students, teens who undoubtedly value interactions with cultures outside their own. They would be expected to value learning a second language more than the typical U.S. high school student does, but U.S. teens' tepid enthusiasm toward learning a second language may also reflect the views of the larger society. Polls indicate that adult Americans are inclined to believe learning a second language is valuable but not essential, unless the second language is English and the learner is an immigrant whose mother tongue is a language other than English.[14]

Differences in educational policies also reinforce these predilections. Learning a second language, for example, is required in 20 European countries—some require learning two foreign languages—and learning a second language begins in the primary grades.[15] In the U.S., on the other hand, taking a second language is not required. Even at the college level, enrollments in languages other than English have experienced declines. A 2013 survey by the Modern Language Association showed the largest decline (6.7% from 2009–2013) for any interval since the survey was first given in 1958.[16]

Conclusion

The 2016 survey of foreign exchange students revealed many of the same findings that emerged from the 2001 survey. Students abroad rarely work at part-time jobs while attending high school; they enjoy spending time with friends, and much of their social lives has migrated to social media; they believe students in their home countries devote more time to schoolwork than American students do; they regard U.S. classes as easier than those at their home schools; they believe that, compared to students back home, U.S. students place a greater importance on success at sports; and they believe success at mathematics or learning a second language does not rise to the same level of importance among their American friends as with peers in their home countries.

It is striking that three of the most lopsided response frequencies from 2001—indicating the pre-eminent role of sports, less time devoted to schoolwork, and relative easiness of classes in the U.S.—are even more pronounced in the 2016 data. The past two decades of education reform in the U.S. have focused on ratcheting up expectations through standards and testing and holding schools accountable for academic progress. Whatever their impact on learning, these efforts appear not to have dramatically altered the impression that U.S. schools, when compared to those of other countries, do not fully embrace inculcating knowledge as the high school's primary institutional mission. Socialization and the production of "well-rounded" citizens are also important objectives of U.S. schooling.

Amanda Ripley's 2013 book, "The Smartest Kids in the World," follows three students as they go overseas to attend high schools in Korea, Poland, and Finland. Ripley also surveyed exchange students, both students from abroad attending U.S. schools and U.S. students going abroad, and drew her survey sample from AFS participants.[17] Two of the survey's questions were replicated from the Brown Center's 2001 surveys. The responses, collected in 2012, mirror the findings reported here.

On the question of whether classes in the U.S. are easier or harder, two-thirds of international students (67%) said classes are much easier in the U.S., and another 25% called them "a little easier." U.S.

U.S. teens' tepid enthusiasm toward learning a second language may reflect the views of the larger society.

The institutional structure of high schools, which is the responsibility of adults, cultivates the adolescent society and is slower to change.

students who had gone to other countries as exchange students reaffirmed this assessment, with 32% rating U.S. classes as "much easier" and 38% labeling them "a little easier." On the importance of sports, 69% of international students said doing well in sports was "much more" important among U.S. students than students in other countries; 22% said "a little more." Of the U.S. students who had gone to other countries as exchange students, 43% said success at sports was "much more" important to their friends in the U.S. and 19% said a "little more important." These findings are persistent.

Teen culture in the U.S. reflects the dispositions of teenagers worldwide, that is true, but it also reflects American dispositions. Whether it is James Coleman's study from the 1950s, the Brown Center Report's studies of high school culture in the 2000s, or Amanda Ripley's account of American students venturing abroad in the 2010s, adults are not the main actors—but their complicity in the story should not be discounted. Kids listen to their parents and teachers, even if it does not always appear so. They also notice the values embraced by the adult world that they are soon to join. Policies championing education reform come and go, but the institutional structure of high schools, which is the responsibility of adults, cultivates the adolescent society and is slower to change. The concluding sentences to "Adolescent Society" remain valuable: "If secondary education is to be successful, it must successfully compete with cars and sports and social activities for the adolescents' attention. … It is up to the adult society to so structure secondary education that it captures this energy."

Part III RACE AND SCHOOL SUSPENSIONS

SCHOOL DISCIPLINE IS A BALANCING ACT. WHETHER DECIDING school policy for an upcoming year or debating the consequences for an individual student arising from a single incident of misbehavior, educators juggle several factors: striking a balance between an orderly campus and a welcoming climate conducive to learning; protecting the safety of all students while recognizing the rights of individuals; treating students equitably but, when warranted, considering individual circumstances that influence behavior; and, in concert with every school's educative mission, convincing students who are behaving badly to correct their behavior while also standing ready to banish anyone who interferes with the learning of others. Getting discipline right is an integral characteristic of a good school. Getting it wrong can be a disaster.

The U.S. Department of Education's Office for Civil Rights caused a stir in 2014 when it released data showing that black students are suspended and expelled at three times the rate of white students. Two years earlier, a report from the Civil Rights Project at UCLA documented large racial disparities in California school districts' disciplinary practices. The report noted that black students are disproportionately dealt the harshest exclusionary penalties—expulsions and out-of-school suspensions.[18] In 2014, the California state legislature passed a state law (AB420) prohibiting public schools from expelling any student or suspending students in third grade or earlier grades for the offense of "willful defiance"—a catchall category of offenses (including disruption) ranging from shouting obscenities at a teacher to forgetting to bring a pencil to class. At the time of AB420's signing by Governor Jerry Brown, willful defiance was the most common offense for out-of-school suspensions, particularly for minority students.[19] However, its use was already in decline, as several districts, including Los Angeles and San Francisco, had previously taken steps to limit expulsions and out-of-school suspensions, regardless of the grade, for willful defiance.[20]

In 2015, just a few days after AB420 took effect, the California Department of Education issued a press release documenting sharp declines in suspensions during the previous two school years.[21] The

department had been providing workshops and other resources on alternatives to out-of-school suspensions, with restorative justice and positive behavioral support programs leading the list.

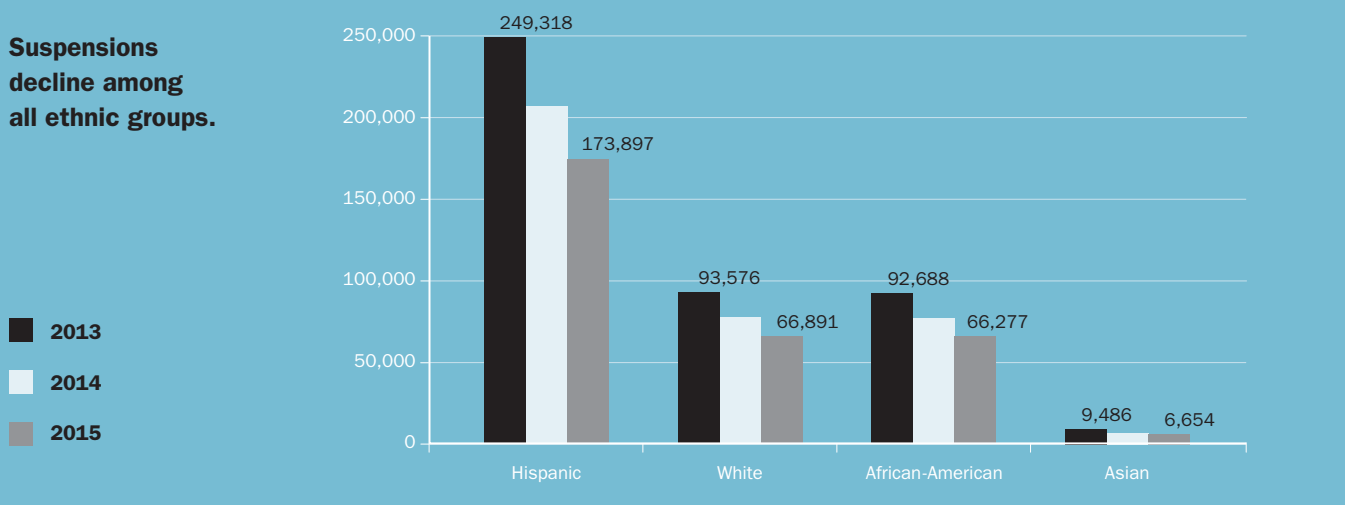
California schools' disciplinary policies are in transition. That offers a unique opportunity to examine the relationship between race and discipline in K-12 education. The study below is a correlational study. It takes no position on out-of-school suspensions or the interventions proposed as alternatives. The data cannot test causal hypotheses or determine whether students' suspensions are just or unjust. The study examines how California schools have changed their practice in regards to out-of-school suspensions and investigates the characteristics of schools associated with both high and low suspension rates of black students.

State trends

From 2013 to 2015, out-of-school suspensions declined dramatically in California. Figure 3-1 displays statistics for the state's four largest ethnic groups, showing the number of suspensions for each group. In this case, a suspension is an event leading to the out-of-school suspension of a student, and even though the student may have committed multiple infractions, the event is counted as one incident. Suspensions of Hispanic students fell from 249,318 in 2013 to 173,897 in 2015, a decline of 30.3%. Suspensions of white (-28.5%), black (-28.5%), and Asian (-29.9%) students declined by similar amounts. Note that the chart begins with 2012 data. The statewide decline in suspensions began a year earlier for all four groups, falling by 12%-13% from 2011 to 2012.

Out-of-school suspensions in California, by ethnicity (2013–2015)

FIGURE 3-1



African-Americans stand out as disproportionately receiving suspensions.

Disparities associated with race or ethnicity cannot be gleaned from simple counts of suspensions. Figure 3-2 reports the ratio of suspensions to enrollment. In 2015, Hispanic students made up over half of California's K-12 enrollment, with 3.3 million students (53.6%). White students constituted 24.6% of enrollment (1.5 million), followed by Asian students, 8.8% (about 550,000 students), and black students, 6.0% (about 370,000 students).[22]

African-Americans stand out as disproportionately receiving suspensions. The 2013 rate of 0.235 means that for every 1,000 black students enrolled in California schools, black students received 235 out-of-school suspensions. That rate dropped to 0.178 in 2015, a decrease of 24.3%. Despite the decrease, the 2015 suspension rate for black students remained much higher than for other ethnic groups. The other groups also experienced declining suspension rates. The 2015 rates for Hispanic (0.052) and white (0.044) students were comparable—one-third to one-fourth the rate for black students—and close to the state average for all students (0.54). The rate for Asians (0.012) was about one-fourth that of Hispanics and whites.

In sum, the black suspension rate stands out on the high side, the Asian rate stands out on the low side, and Hispanic and white rates are near the state average. The number of out-of-school suspensions fell dramatically between 2013 and 2015, but black students were still disproportionately suspended compared to other ethnic groups in 2015.

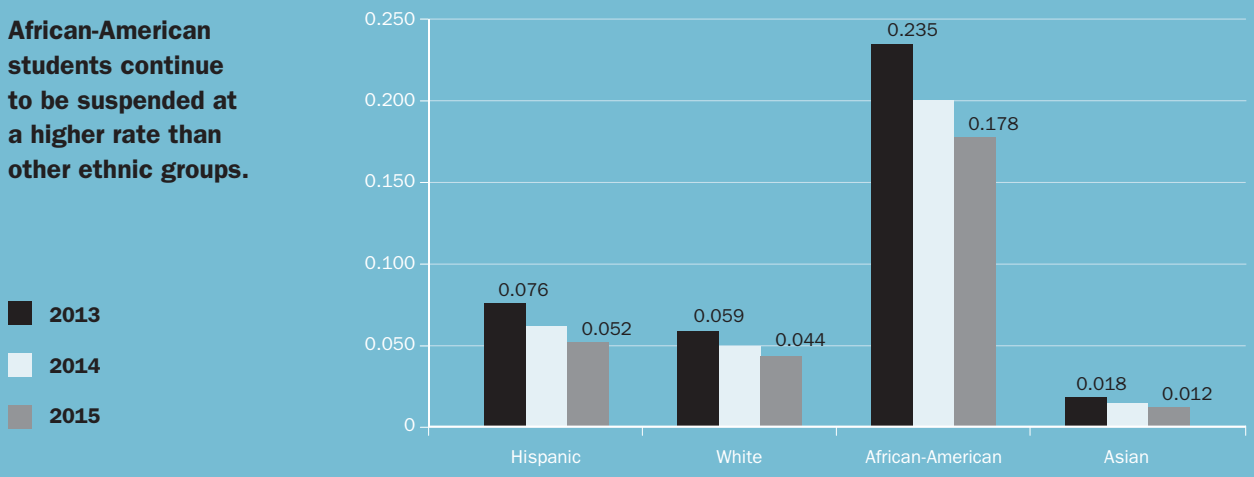
School-level analysis of African-American suspension rates

Let's dig deeper into the disproportionate number of suspensions received by black students and see how schools with high

FIGURE 3-2

Suspension rates in California, by ethnicity (2013–2015)

African-American students continue to be suspended at a higher rate than other ethnic groups.



suspension rates for black students differ from schools with low rates. School-level data were downloaded from the California Department of Education, California Longitudinal Pupil Achievement Data System (CALPADS) website.[23] The following were dropped: schools with total enrollment of fewer than 50 students, special education schools, continuation high schools, schools in juvenile delinquency facilities, alternative schools, community day schools, and schools lacking a full panel of data for 2013–2015. A total of 7,180 schools remained available for analysis after the screens were applied. In 2015, these schools served about 5.4 million of California's 6.2 million K-12 students (approximately 87%).

In the discussion below, academic years are labeled by their June calendar year (e.g., the 2012–2013 academic year is "2013"). The analysis is limited to out-of-school suspensions, hereafter simply called "suspensions" to simplify wording. Suspensions that are deployed in school, which send a misbehaving student to a supervised "time-out" spot on campus, are less exclusionary than—and normally considered an alternative to—out-of-school suspensions. Each incident leading to a suspension is counted as one observation, even though a student may have been charged with more than one infraction during the incident. Suspensions occurring at different points in time and arising from different incidents are counted separately even though they may involve the same student. Note that other studies of school discipline may focus on students instead of suspensions and use what is known as "unduplicated counts" as the analytical unit. In those studies, for example, a student who has been suspended six times during a school year will count as one student, but in the analysis below, those six suspensions would all count towards a school's total number of suspensions.[24]

Table 3-1 sorts schools into two groups, those with high suspension rates (0.05 or greater) for African-American students and those with low rates (less than 0.05). A school's suspension rate was calculated as the number of suspensions of black students divided by the number of black students enrolled at the school. The 0.05 suspension rate was used as the dividing line because it is the state average for all students of all races. Put in plain English, the definition of a high-suspension-rate school is a school that has reported five or more suspensions of black students for

Suspension rates for African-American students, by school (2013–2015) TABLE 3-1

African-American suspension rate	2013 no. of schools	2013 suspensions	2015 no. of schools	2015 suspensions	Change in suspensions
High (5% or greater)	2,177 (30.3%)	51,795	1,930 (26.9%)	35,424	-16,371
Low (less than 5%)	3,357 (46.8%)	202	3,546 (49.4%)	139	-63
Asterisk (unknown)	1,646 (22.9%)	*	1,704 (23.7%)	*	*
Total	7,180	51,997	7,180	35,563	-16,434

Note: In 2013, 3,240 of Low schools reported zero African-American suspensions. In 2015, Low schools with zero African-American suspensions totaled 3,432. Schools reporting 1–10 suspensions are recorded with "*" in state data bases to protect student confidentiality. Total number of suspensions refers to known suspensions.

Middle school appears to be the chronological dividing line for when African-American suspension rates escalate.

every 100 black students enrolled. The high-suspension-rate schools totaled 2,177 schools in 2013; 3,357 schools had a suspension rate below 0.05 and are designated low-suspension-rate schools. Most of the low-suspension-rate schools (3,240), in fact, did not suspend any black students in 2013.

The contrast between the two groups is shocking. The high-suspension-rate schools, although the smaller group, suspended 51,795 African-American students in 2013 and 35,424 in 2015. The low-suspension-rate schools suspended only 202 black students in 2013 and 139 in 2015. The high-suspension-rate schools are where the most dramatic decline in the number of suspensions has occurred, but they are also where most suspensions of black students continue to occur. The high-suspension-rate schools enrolled about 130,000 black students in 2015, making their suspension rate (0.27) almost 160 times larger than the low-suspension-rate schools; their rate was approximately 0.0017 (black enrollment of approximately 80,000 students).

The third row, designated as "asterisk" schools, are schools that suspended between one and 10 black students, but the exact number is unknown. For the sake of shielding the identity of students, California's public databases report only a "*," and not the number of suspensions, when the annual count is fewer than 11 suspensions. Despite that data constraint, many schools can still be identified as a high- or low-rate suspension school based on the number of African-American students enrolled at the school. An "asterisk" school with 20 or fewer black students was classified as a high-suspension-rate school (even the lowest possible count, one, puts them at a 0.05 rate), and a school with 201 or more black students was classified as a low-suspension-rate school (even with the highest possible count, 10, the rate is less than 0.05). In the tables below, statistics for the remaining "asterisk" schools (about 23.7% of the analytical sample in 2015) are reported. The "asterisk" schools remain in the analysis, but their individual suspension rates for black students cannot be determined from available data.

In sum, the analysis focuses on two groups of schools, one numbering nearly 2,000 and the other about 3,500, with profoundly different African-American suspension incidence rates. Let's examine some other characteristics that differentiate these schools.

Suspension rates for African-American students, by grade configuration of school (2015) — **TABLE 3-2**

African-American suspension rate	Elementary		K-8		Middle school		High school	
	Count	%	Count	%	Count	%	Count	%
High (5% or greater)	685	18.9%	177	26.7%	530	42.9%	409	38.5%
Low (less than 5%)	2098	57.9%	373	56.3%	389	31.5%	317	29.9%
Asterisk (unknown)	842	23.2%	112	16.9%	315	25.5%	335	31.6%
Total	3,625		662		1,234		1,061	

Note: Percentages shown are of column totals. Schools with other grade configurations are not shown (n = 598).
The shaded cells indicate whether low or high suspension rates are greater.

Grade configuration

Table 3-2 exhibits the grade configuration of schools. The shaded cells indicate whether low or high suspension rates for black students are dominant. Schools with younger children tend to have low suspension rates. The ratio of low (57.9%) to high (18.9%) in elementary schools is about 3-to-1. In K-8 schools, it is about 2-to-1—56.3% to 26.7%. The pattern reverses in middle schools and high schools.

Middle schools have the largest percentage of high-suspension-rate schools, 42.9%, surpassing the 31.5% of schools with a low suspension rate. Among high schools, 38.5% have high suspension rates, compared to 29.9% with low suspension rates. Middle school appears to be the chronological dividing line for when African-American suspension rates escalate. Elevated rates are also prevalent in high school, but not quite as pronounced as for middle schools. For both types of secondary schools, suspension rates declined from 2013 to 2015 (not shown in the table). Middle schools with high suspension rates were 50.9% of the middle school sample in 2013; for high schools, the figure was 41.5%.

School size

Large schools present management challenges that administrators in small schools do not face. A lot of kids sharing the same physical space can lead to problems. Researchers have long suspected that small schools are better able to cultivate personal connections between adults and students that promote positive social behaviors. Research suggests larger schools may face more serious behavior problems, but the literature is not definitive.[25]

Table 3-3 shows African-American suspension rates in California schools of three different sizes. More than three-quarters (78.0%) of small schools, those with fewer than 200 students, have low suspension rates, while only 16.7% have high suspension rates. The ratio is almost 5-to-1 favoring schools with low suspension rates. The middle size group in the table, schools with 201–1,300 students, are also more likely to evidence low suspension rates, with about twice as many lows as highs (51.3% v. 26.0%). Large schools stand out, with more schools identified as high-suspension-rate schools (38.2%) than low (22.4%). There are two caveats, however. School size is correlated with grade levels—middle and high

Suspension rates for African-American students, by school size (2015) — TABLE 3-3

African-American suspension rate	Student population 0–200		Student population 201–1,300		Student population 1,300+	
	Count	%	Count	%	Count	%
High (5% or greater)	53	16.7%	1,587	26.0%	290	38.2%
Low (less than 5%)	248	78.0%	3,128	51.3%	170	22.4%
Asterisk (unknown)	17	5.3%	1,387	22.7%	300	39.5%
Total	318		6,102		760	

Note: School size sample mean = 749, SD = 543.
The shaded cells indicate whether low or high suspension rates are greater.

Schools in wealthier communities are less likely to suspend African-American students than other schools.

Suspension rates for African-American students, by percentage of students qualifying for free and reduced-price meals (FRPM) (2015)

TABLE 3-4

African-American suspension rate	FRPM 0–33%		FRPM 33–89%		FRPM>89%	
	Count	%	Count	%	Count	%
High (5% or greater)	246	16.3%	1,265	29.2%	419	31.2%
Low (less than 5%)	972	64.4%	1,962	45.4%	612	45.5%
Asterisk (unknown)	292	19.3%	1,099	25.4%	313	23.3%
Total	1,570		4,326		1,344	

Note: FRPM sample mean = 61.2%, SD = 28.2%

The shaded cells indicate whether low or high suspension rates are greater.

schools are typically larger than elementary schools. As we just saw, schools serving adolescents are prone to more suspensions than schools with younger children. In addition, the large number of "asterisk" schools muddies the water in terms of schools with more than 1,300 students.

Free and reduced-price meals

Table 3-4 presents suspension data broken out by the percentage of students qualifying for free and reduced-price meals (FRPM), a traditional indicator of school poverty. The first column, FRPM 0–33%, represents schools in wealthier communities; the far right column, FRPM>89%, represents high-poverty schools; and the middle column represents schools that are closer to the state average on the indicator, ranging from one standard deviation above to one standard deviation below the mean.

Schools in wealthier communities are less likely to suspend African-American students than other schools. Almost two-thirds (64.4%) suspend black students at a low rate, whereas only 16.3% suspend at a high rate. The proportion of high-suspension-rate schools increases in tandem with increased poverty. In the middle of the FRPM distribution, 29.2% of schools qualify as high-suspension-rate schools. Surprisingly, though, the pattern does not continue linearly. Schools serving the poorest students, in which 89% or more qualify for FRPM, exhibit similar suspension rates as schools near the FRPM mean: 45.5% are low-suspension-rate schools and 31.2% have high suspension rates. The proportion of students in poverty at a school appears to be associated with black suspension rates, but it is not as powerful a correlate as school size.

Again, there are caveats. Free and reduced-price meals (FRPM) is an imperfect proxy for school poverty. The federal meals program allows families with incomes up to 185% of the official poverty level to qualify for reduced-price meals and families up to 130% of the poverty level to receive free meals. In addition, under the Community Eligibility Provision, non-poor students may receive FRPM if they attend schools in which a majority of students are poor.[26] FRPM is still the most widely available school-level indicator of poverty, but it is a noisy measure—and getting noisier.

Suspension rates for African-American students, by percentage of African-American enrollment (2015)

TABLE 3-5

African-American suspension rate	African-American enrollment 0–6%		African-American enrollment 6–16%		African-American enrollment >16%	
	Count	%	Count	%	Count	%
High (5% or greater)	1,191	23.2%	402	29.9%	337	47.5%
Low (less than 5%)	3,189	62.2%	227	16.9%	130	18.3%
Asterisk (unknown)	745	14.5%	716	53.2%	243	34.2%
Total	5,125		1,345		710	

Note: Percentage black enrollment sample mean = 6.0%, SD = 9.4%

The shaded cells indicate whether low or high suspension rates are greater.

Black proportion of school enrollment

The harmful effects of racial segregation in public schools are well documented. In particular, black students who attend racially isolated schools are more likely to experience lower academic achievement, exposure to gangs, and disruptions during classroom instruction. One study labeled this the "Racial School Climate Gap."[27] A 2014 multilevel analysis of suspension data found the percentage of black enrollment to be the strongest school-level predictor of out-of-school suspensions. The analytical model also included student-level variables, including race, poverty status, and the severity of infractions. Surprisingly, attending a school with a high proportion of black enrollment only contributed slightly less to a student's probability of being suspended as committing the infraction of fighting or battery.[28]

Table 3-5 displays suspension rates disaggregated by the proportion of black students enrolled in schools. Schools with the lowest percentage of black enrollment (0%-6%) do not suspend black students as frequently as schools with larger black populations. Almost two-thirds (62.2%) are low-suspension-rate schools, yet nearly half of the schools with black enrollment larger than 16% are high-suspension-rate schools (47.5%).

Why is this? One explanation is related to segregation. Schools with large black populations may be located in unsafe neighborhoods and consequently adopt tighter disciplinary policies (e.g., "zero tolerance"), with rules enforced by metal detectors, hallway cameras, and security guards. However, the data in Table 3-5 suggest that other factors are at work. A nearly 3-to-1 ratio favoring low-suspension-rate schools flips to a nearly 3-to-1 ratio favoring high-suspension-rate schools, and that reversal becomes apparent with schools in which African-American students are 16% or more of enrollment. Sixteen percent is not racial isolation.

Racial bias or stereotyping cannot be ruled out in either the development or implementation of school rules. Beginning in middle school, African-American students are more likely than Asian and white students to say they are treated unfairly when it comes to school discipline.[29] Black

Nearly half of the schools with black enrollment larger than 16% are high-suspension-rate schools.

students are also more likely to come from family backgrounds associated with school behavior problems; for example, children ages 12–17 that come from single-parent families are at least twice as likely to be suspended as children from two-parent families. Almost two-thirds of black children in California live in single-parent households.[30]

Discussion

State officials in California have been urging schools to reduce out-of-school suspensions. The number of out-of-school suspensions fell by more than 30% from 2013 to 2015. Declines of approximately equal magnitude were registered among students in all four of the state's major racial/ethnic groups—Hispanic, white, Asian, and African-American. One objective of discipline reform has not been met: Black students continue to be suspended at rates disproportionate to their share of the student population.

What can state data tell us about this disproportionality? The current study examined suspension data from 7,180 schools. The schools were divided into two groups, signifying high- and low-suspension rates, based on whether black student suspensions fell above or below 5% of black student enrollment. High-suspension-rate schools numbered 2,177 in 2015. They issued more than 35,000 suspensions of black students. The group of low-suspension-rate schools was much larger at 3,546 schools, but only 139 of their suspensions involved black students. These differences are stunning. How else do the schools differ? The analysis focused on two structural characteristics—grade configuration and size—and two demographic characteristics—the percentage of students qualifying for free and reduced-price meals and the proportion of African-American student enrollment.

Middle schools and high schools are more likely to have high African-American suspension rates; elementary schools and K-8 schools tend to have low rates. Large schools have higher suspension rates for black students than small schools. Schools with more students qualifying for free and reduced-price meals, a proxy for poverty, have higher African-American suspension rates than schools in wealthier communities. Schools enrolling a greater percentage of black students exhibit higher suspension rates for black students than schools with fewer black students.

The study's data cannot definitively confirm that the decline in suspensions came about because of state initiatives promoting discipline reform. UCLA's Civil Rights Project calculated that 77% of the decline in state suspensions from 2012–2014 were in the disruption/defiance category.[31] The decline in suspensions for disruption/defiance is certainly in accord with state preferences, especially as encoded in state legislation, AB420, but as pointed out above, several districts had already taken their own steps to limit defiance suspensions before the law took effect. Like other policies, upper-level (state or federal) policy initiatives may ratify trends already underway rather than start new ones. Reducing suspensions—and the disparate impact of suspensions on black students—has been the subject of national attention and discussion for several years.

Alternatives to suspensions

Suspended students may miss several days of schools and have long periods of time without adult supervision. Out-of-school suspensions are associated with low achievement, poor attendance, and

juvenile crime—a combination of undesirable outcomes that can push students into what has been called the "school-to-prison pipeline." Restorative justice has become a popular alternative to out-of-school suspensions, typically featuring a meeting of victims, perpetrators, parents, teachers, administrators, and a counselor or psychologist. The goal is to get misbehaving students to take responsibility for their behaviors and the consequences that others have suffered. Evaluations of the effectiveness of the intervention are scant. A 2016 study of longitudinal data from Denver Public Schools found restorative practices offered to students in one semester were associated with reductions in suspensions during the following semester. The authors noted that no randomized control trial (RCT) of the program has been conducted to date.[32] A 2016 research review from WestEd identified three RCTs of restorative interventions underway, with the earliest evaluation expected to be completed in 2018.[33]

Studies of disciplinary reform tend to focus on suspended students. What about those students' peers?

Research suggests that misbehaving students take a toll on the education of others. A 2014 report from Ofsted, the United Kingdom inspectorate of schools, estimated that each year, British teachers lose the equivalent of 38 days of instruction dealing with even low-level misbehavior. Carrell, Hoekstra, and Kuka (2016) investigated the long-term impact of classroom disruption on peers by linking Florida elementary school data to later outcomes. Exposure to a single disruptive peer in an elementary class of 25 students leads to a 1.6% reduction in college enrollment, a 2.6% reduction in the likelihood of receiving a college degree (including two-year degrees), and a 3–4% loss in earnings.

Discipline reform will rise or fall on whether disruptive behaviors, not just suspensions, are diminishing. Around the country, teachers have challenged the merits of discipline reform by warning about deteriorating safety and order at schools. In December 2016, The Fresno Bee reported that 70 of 85 teachers at McLane High School, a school publicly praised for its restorative justice program, signed a petition demanding stricter, more consistent discipline at the school.[34] The petition claims the campus has experienced constant disruption and fighting—and that teachers have been both physically and verbally assaulted.

> "This is unfair to students who come to school ready to learn and the teachers ready to engage them," the petition states. "Our students are losing valuable classroom instructional time, which is depriving them of their right to an education, decreasing their chances at success in college and careers, and affecting the morale of staff members."

Consider the data above on racial composition of schools. It is true that black students are suspended at disproportionately high rates, but the negative effects of disruptive students on rule-abiding peers almost certainly fall disproportionately on black students as well.

Restorative justice has become a popular alternative to out-of-school suspensions.

Policymakers need to consider a broader array of possible solutions to the problem of disciplinary equity.

Conclusion

The foregoing analysis suggests that policymakers should consider altering the organizational characteristics of schools as a strategy for reducing disparities in black suspensions.

It will not be easy. Racial or socio-economic desegregation of schools has proven difficult since the 1970s. Breaking large schools up into smaller units may reduce suspension rates for all students—and especially for black students—but that means reassigning students to new schools. Researchers have questioned the merits of isolating middle school students in their own grade configuration for years. A North Carolina study found that students who attended sixth grade at a middle school were more likely to be suspended, and later to repeat a grade or drop out of school, compared to counterparts who attended sixth grade in an elementary school.[35] A longitudinal study of New York City sixth graders found that attending sixth grade in a middle school, as opposed to a K-8 school, produced a negative impact on achievement that began in the first year and extended throughout the middle school years. In surveys of both students and parents, respondents described middle school campuses as less safe, peer behavior as more immature and antisocial, and the overall quality of education as lower than in K-8 schools.[36]

Restorative justice interventions or other discipline reforms may prove to be successful alternatives to out-of-school suspensions; however, their efficacy has not yet been evaluated in rigorously designed evaluations. Suspensions are declining, but reforms have also been unpopular with some teachers and principals. Policymakers need to consider a broader array of possible solutions to the problem of disciplinary equity and take policy guidance from studies that can assess causal impact. Those evaluations should include staff morale, the effect on learning environments, and the learning of rule-abiding peers as measured outcomes. Breaking up large schools to create smaller campuses and reconfiguring middle schools as K-8 schools, at least based on the correlational evidence presented here, appear to be structural changes with the potential to support disciplinary equity.

NOTES

1 Christine Armario, "'Wake-up call': U.S. students trail global leaders," Associated Press, December 7, 2010.

2 For scores from the First International Math Study (FIMS), see the 2010 Brown Center Report on American Education, "Part I: International Tests," Table 1-3, p. 9.

3 National Center for Education Statistics (NCES), "Selected Findings from TIMSS 2015." Accessed at: https://nces.ed.gov/timss/timss2015/findings.asp

4 See the 2008 Brown Center Report on American Education, "Part I: The Use and Misuse of International Assessments," pp. 7–18, and the 2012 Brown Center Report on American Education, "Part III: Misinterpreting International Test Scores," pp. 24–31. Also see, "International Test Are Not All the Same." Accessed at: https://www.brookings.edu/research/international-tests-are-not-all-the-same/.

5 National Center for Education Statistics (NCES), "Comparing NAEP, TIMSS, and PISA in Mathematics and Science." Accessed at: https://nces.ed.gov/timss/pdf/naep_timss_pisa_comp.pdf

6 Eric A. Hanushek and Ludger Woessmann, "The Economics of International Differences in Educational Achievement," *Handbook in Economics, Volume 3: Economics of Education,* pp. 89–200 (San Diego: North-Holland, 2011).

7 Heiner Rindermann, "The g-factor of International Cognitive Ability Comparisons: The Homogeneity of Results in PISA, TIMSS, PIRLS, and IQ-Tests Across Nations," *European Journal of Personality*, *Volume 21, Issue 5,* pp. 667–706. (John Wiley & Sons, 2007)

8 Eckhard Klieme. "TIMSS 2015 and PISA 2015: How Are They Related on the Country Level?" German Institute for International Educational Research (2016). Accessed at: http://www.dipf.de/de/publikationen/pdf-publikationen/Klieme_TIMSS2015andPISA2015.pdf

9 Angel Gurría, "Presentation of the PISA 2010 Results." (OECD: 2010). Accessed at: https://www.oecd.org/unitedstates/presentationofthepisa2010results.htm

10 James S. Colman, *The Adolescent Society* (Oxford: Free Press of Glencoe, 1961), p. 3.

11 Common Sense Media, *Children, Teens, and Reading: A Common Sense Media Research Brief (2014)*.

12 NAEP Data Explorer. Accessed at: http://nces.ed.gov/nationsreportcard/naepdata

13 Bureau of Labor Statistics summary of regulations. Accessed at: https://www.dol.gov/whd/state/nonfarm.htm

14 Chris McComb, "About One in Four Americans Can Hold a Conversation in a Second Language," Gallup (2001). Accessed at: http://www.gallup.com/poll/1825/about-one-four-americans-can-hold-conversation-second-language.aspx

15 Kat Devlin, "Learning a foreign language a 'must' in Europe, not so in America," Pew Research Center (2015). Accessed at: http://www.pewresearch.org/fact-tank/2015/07/13/learning-a-foreign-language-a-must-in-europe-not-so-in-america/

16 David Goldberg, Dennis Looney, and Natalia Lusin, *Enrollments in Languages other than English in United States Institutions of Higher Education, Fall 2013* (New York: Modern Language Association, 2015).

17 For the 2012 survey results, see Appendix II in *The Smartest Kids in the World*, by Amanda Ripley (New York: Simon & Shuster, 2013).

18 Daniel J. Losen, Tia Martinez, and Jon Gillespie, *Suspended Education in California* (UCLA: The Civil Rights Project, 2012).

19 The governor had vetoed two earlier versions of the bill that banned willful defiance suspensions in all schools, including middle and high schools. The states two teachers unions opposed the earlier bills but took a neutral position on AB420.

20 Susan Frey, "New Law Limits Student Discipline Measure," *EdSource*, September 28, 2014.

21 California State Department of Education, "State Schools Chief Torlakson Reports Significant Drops in Suspensions and Expulsions for Second Year in a Row," January 14, 2015. Accessed at: http://www.cde.ca.gov/nr/ne/yr15/yr15rel5.asp. Also see similar press release from 2016: http://www.cde.ca.gov/nr/ne/yr16/yr16rel5.asp

22 California has five other reporting categories for race and ethnicity: Flipino, Not Hispanic (2.5%), Two or More Races (2.4%), American Indian or Alaskan Native (0.7%), Pacific Islander, Not Hispanic (0.6%), and None Reported (0.7%). In this paper, Hispanic or Latino of Any Race is referred to as "Hispanic," Asian, Not Hispanic as "Asian," African American, Not Hispanic as "African American" or "Black," and White, Not Hispanic as "White."

23 Data were combined from the following files for each year of analysis: Free and Reduced Price Meal Data, Enrollment Data: Enrollment by School, and Expulsion and Suspension Data. Except for suspension data, which are collected at the end of the year, all other data were collected on census day, the first Wednesday in October of the school year.

24 The National Cancer Institute at the National Institutes of Health explains how cancer incidence rates are computed similarly: "The numerator of the incidence rate is the number of new cancers; the denominator is the size of the population. The number of new cancers may include multiple primary cancers occurring in one patient." Accessed at: https://seer.cancer.gov/statistics/types/incidence.html

25 Several papers on the effects of school and class size can be found in: "Brookings Papers on Education Policy, 2006/2007," eds. Tom Loveless and F.M. Hess, (Washington, DC: Brookings Institution Press, 2007).

26 Schools and districts qualify by using data from other means-tested programs to establish extensive poverty in the community. Accessed at: https://www.fns.usda.gov/school-meals/community-eligibility-provision

27 Adam Voight. *The Racial School-Climate Gap*. (San Francisco: WestEd, 2013).

28 Russell J. Skiba, Choong-Geun Chung, Megan Trachok, Timberly L. Baker, Adam Sheya, and Robin L. Hughes, "Parsing Disciplinary Disproportionality: Contributions of Infraction, Student, and School Characteristics to Out-of-School Suspension and Expulsion," *American Education Research Journal*, August 2014, 52(4): pp. 640–670.

29 Adam Voight. *The Racial School-Climate Gap*. (San Francisco: WestEd, 2013), p. 4.

30 Kids Count Data Center, "Children in single-parent families by race," 2015. Accessed at: http://datacenter.kidscount.org/data/tables/107-children-in-single-parent-families-by#detailed/2/2-52/false/573,869,36,868,867/10,11,9,12,1,185,13/432,431

31 Daniel J. Losen, Michael A. Keith II, Cheri L. Hodson, Tia E. Martinez, and Shakti Belway, *Closing the School Discipline Gap in California: Signs of Progress* (UCLA: The Civil Rights Project, 2015).

32 A 2015 meta-analysis of 12 randomized control trials (RCT) of restorative interventions in UK and Australian criminal justice settings found reduced repeat offenses on a short–term basis but weaker long term effects. It is unclear whether restorative approaches with adults under the auspices of courts and police are applicable to school discipline.

33 T. Fronius, H.Persson, S. Guckenburg, N. Hurly, and A. Petrosino, *Restorative Justice in U.S. Schools: A Research Review* (San Francisco: WestEd, 2016).

34 Mackenzie Mays, "Restorative justice? Teachers say McLane High classrooms are spiraling out of control," *The Fresno Bee*, December 10, 2016.

35 P.J. Cook, R. MacCoun, C. Muschkin, and J. Vigdor, "Should Sixth Grade Be in Elementary or Middle School? An Analysis of Grade Configuration and Student Behavior," NBER Working Paper #12471 (August 2006).

36 J.E. Rockhoff and B.B. Lockwood, "Stuck in the Middle: How and Why Middle Schools Harm Achievement," *Education Next* (Fall 2010), pp. 69–75.

BROOKINGS
1775 Massachusetts Avenue, NW • Washington, D.C. 20036
Tel: 202–797–6000 • Fax: 202–797–6004
www.brookings.edu

The Brown Center on Education Policy
Tel: 202–797–6090 • Fax: 202–797–2480
www.brookings.edu/brown