What Is the Matter with America's Schools?[1]

ROBERT M. HAUSER
Executive Officer
American Philosophical Society

For almost four decades, politicians, educational leaders, and researchers have complained about America's standing in international comparisons of academic achievement and thus cast broad criticism of the quality of U.S. primary and secondary schooling. Such complaints ignore and fail to address the source of America's poor international standing. Only one thing is the matter with academic achievement in America's primary and secondary schools: the achievement gap between non-Hispanic whites or Asians and the nation's two largest minority groups—non-Hispanic Blacks and Hispanics. Relative to other nations, the academic achievement of U.S. non-Hispanic whites and U.S. Asians is impressive. The fault here does not lie with minority youth, for we have the knowledge to address the achievement gap. The problem is political, not technical. By addressing the achievement gap in a serious and sustained way, our political and educational leaders could establish the United States as an international leader in academic achievement.

First, I show that U.S. non-Hispanic whites and U.S. Asians have high test scores in reading, science, and mathematics relative to the highest scoring nations, both on the Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS). That is, there is no global problem with American primary and secondary schooling. The problem is that U.S. non-Hispanic Blacks and U.S. Hispanics have lower average scores on these assessments, and that accounts for the lower overall scores of U.S. students. Further, I show that differences in social and economic background are unlikely to account fully for observed racial-ethnic differentials.

1 This is a revised version of a presentation given at the November 2017 Meeting of the American Philosophical Society and in various academic venues. Several friendly critics have been most helpful. They include Richard T. Campbell, Adam Gamoran, Eric Grodsky, David Grusky, Bethany B. Hauser, Michael Hout, Min-Hsiung Huang, Shu-Ling Tsai, John Robert Warren, and Yu Xie. Special thanks go to Alyn Turner, who has valiantly and energetically urged and helped me to get this right from the beginning. All errors and omissions in the work are strictly my fault.

Second, I ask whether there are existing, large-scale, long-term exemplars of U.S. school systems that substantially reduce score differences of non-Hispanic Blacks and Hispanics relative to non-Hispanic whites and Asians. While there may be multiple exemplars, I focus on schools operated by the Department of Defense Education Activity (DoDEA), where racial-ethnic test-score differences are far smaller than in other public schools. Finally, I discuss reasons for the success of DoDEA schools and consider whether they may be a model for educational reform in the United States.

It is time for U.S. educators and policy makers to recognize that the achievement gap is the main driver of our relatively poor international standing and to mobilize the knowledge and resources to solve this problem. The fault does not lie in the academic potential of minority students. There are well-established models of success for all students. Rather, American educational policies and practices typically fail students in poorly endowed schools where most children are non-Hispanic Black or Hispanic.

A Nation at Risk, Then and Now

Critics have complained continuously and visibly about the quality of America's public primary and secondary schools for the past third of a century, beginning with the publication of *A Nation at Risk*, the report of the National Commission on Excellence in Education (U.S. Department of Education 1983). The key passage in the report, which galvanized public opinion, stated, "If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war. As it stands, we have allowed this to happen to ourselves" (9).

Most of the evidence in the report was based on achievement test scores, for example: "International comparisons of student achievement, completed a decade ago, reveal that on 19 academic tests American students were never first or second and, in comparison with other industrialized nations, were last seven times. . . . The College Board's Scholastic Aptitude Tests (SAT) demonstrate a virtually unbroken decline from 1963 to 1980. Average verbal scores fell over 50 points and average mathematics scores dropped nearly 40 points" (11).

The findings of *A Nation at Risk* were fundamentally flawed. Overall test score averages declined, but this was evidence of an educational success story, rather than of failing schools. Observed negative trends in overall academic achievement test scores were largely a consequence of the changing demographic characteristics of test-takers (Carson, Huelskamp, and Woodall 1993; Huelskamp 1993). At that

time, taking college entrance exams was both voluntary and costly. However, as educational opportunity increased, lower-scoring social and economic groups were increasingly likely to take the SAT. Moreover, during the period when aggregate SAT scores declined, scores actually increased within major social groups. However, the growth of scores within groups was not large enough to counterbalance the rising numbers of test-takers in lower-scoring groups.

Despite these facts, there have been persistent claims of support for the findings in *A Nation at Risk* (Stedman 1994, 1997; Bracey 1995; Stedman, Mullis, and Timpane 1998). As noted above, one of the charges in *A Nation at Risk* was that U.S. students fared poorly in international comparisons. There has been no letup in such criticisms over the past 30 years (e.g., Stedman 1997). In December 2016, shortly after the release of findings from the 2015 round of the Organisation for Economic Co-operation and Development's (OECD) PISA, the U.S. Secretary of Education at the time, John King, Jr., said, "U.S. students are running in place . . . we're losing ground" (Resmovits 2016). The 2015 PISA findings featured world-class science and math performance in four provinces of China, and an education blogger declared, "In 1957, the launch of Sputnik by the Soviet Union forced the United States to respond with an enormous national investment in science and mathematics education. This challenge, from China, seems to this observer to be no less important" (Tucker 2016).

The U.S. Secretary of Education from 2017 to 2021, Betsy DeVos, picked up where King left off in this regard. Devos was quoted at an early 2017 Brookings Institution event: "I'm not sure how they could get a lot worse on a nationwide basis than they are today. I mean, the fact that our PISA scores have continued to deteriorate as compared to the rest of the world and that we've seen stagnant at best results with the NAEP scores over the years. I'm not sure we can deteriorate a whole lot." To this, her interlocutor, Russ Whitehurst, replied, "NAEP scores had in fact gone up significantly over the past 20 years for low-performing students. And U.S. students have never done well on the PISA tests" (Strauss 2017).

Other media accounts of the 2015 PISA findings—in *The New York Times, Washington Post,* and *Los Angeles Times*—led with the middling performance of U.S. students. While there were references to socioeconomic differences in test performance—highlighted in a 2013 report on international comparisons (Carnoy and Rothstein)—no one acknowledged the well-known differentials in performance among racial-ethnic groups in the United States.

Although American race and ethnic differentials in academic achievement are ignored in international comparisons, they always

figure prominently in reports about the premier American tests, those of the National Assessment of Educational Progress (NAEP). By way of example, a *New York Times* report about school funding in Connecticut stated, "On the most recent National Assessment of Educational Progress, 46 percent of white fourth graders across the country read at or above 'proficient,' compared with just 18 percent of their black peers. And 51 percent of white fourth graders were at or above proficient in math, compared with 19 percent of black fourth graders" (Zernike 2016).

This raises obvious questions and suggests that statements of concern and policy proposals miss the point: How does the performance of U.S. racial-ethnic groups affect the international standing of American education? How do U.S. racial-ethnic groups fare in international comparisons? What explains the differences among them? Has academic performance in the large-scale international assessments declined in the United States relative to other nations? Fortunately, it is possible to answer these questions using unpublished tabulations and micro-data from the National Center for Education Statistics (NCES).

U.S. Academic Achievement in International Perspective

Consider the 2015 achievement test scores from PISA. Since 2003, PISA has regularly tested 15-year-olds in 72 countries in reading, science, and math. When nations (and some city-states) are compared on these assessments of reading, science, and math literacy, the United States always ends up in the middle of the pack (Institute of Education Sciences and National Center for Education Statistics 2016), thus leading to regular pronouncements of the mediocrity of American schools.

- In reading literacy, the U.S. average score of 497 ranked 24th in the world and was barely higher than the OECD average. Singapore, Hong Kong, Canada, Finland, and Ireland each scored well above the United States. In all, some 14 nations had average scores that were significantly higher than those in the United States.[2]

- In science literacy, the U.S. average score of 496 ranked 25th and, again, was barely higher than the OECD average. Singapore, Japan, Estonia, Taiwan, and Finland each scored much

2 In the U.S. data from PISA 2015, the standard deviation of reading literacy was 95.5, so the U.S. mean of 497 was 0.40 standard deviations below that of top-ranked Singapore, where the mean score was 535.

higher than the United States.[3] Overall, 18 nations scored significantly higher than the United States.[4]

- In math literacy, U.S. performance was even worse. The U.S. average of 470 ranked 40th, significantly below the OECD average of 490. Thirty-six nations—more than half of those participating—scored significantly higher than the United States. The top seven positions were held by Asia: Singapore, Hong Kong, Macau, Taiwan, Japan, four provinces of mainland China, and the Republic of Korea.[5, 6]

Racial-Ethnic Differentials in Academic Achievement

Multiple sources of international achievement test data, including TIMSS and PISA, demonstrate that in reading, science, and math, the mediocre aggregate performance of the United States is primarily attributable to the relatively low performance of Hispanic and Black students. U.S. non-Hispanic whites and U.S. Asians consistently score well in comparison with leading nations. These international comparisons show that America's elementary and secondary schools are not failing overall. Rather, they fail to serve large numbers of minority students, who comprise a relatively large share of the student population. To be sure, the data presented here refer to national samples and thus, necessarily, ignore differences in educational quality and achievement in every population group and at every level from the classroom to the state (Hanushek, Peterson, and Woessmann 2010; Peterson et al. 2011; Hanushek, Peterson, and Woessmann 2012). Moreover, other nations may also have minority populations, e.g., recent immigrants, whose academic performance may affect their international standing.

Reading Literacy

How does the racial-ethnic composition of American students affect international findings? Table 1 shows how PISA participants scored and ranked in reading literacy, but differs from the NCES presentation

3 For political reasons, PISA lists Taiwan as Taipei.

4 In the U.S. data from PISA 2015, the standard deviation of science literacy was 95.5, so the U.S. mean of 496 was 0.63 standard deviations below that of top-ranked Singapore, where the mean score was 556.

5 In the U.S. data from PISA 2015, the standard deviation of math literacy was 83.8, so the U.S. mean of 470 was 1.12 standard deviations below that of top-ranked Singapore, where the mean score was 564.

6 For political reasons, PISA lists Taiwan as Taipei.

by treating U.S. whites, U.S. non-Hispanic Blacks, U.S. Hispanics, and U.S. Asians as separate groups.[7] There are now 74 instead of 70 ranked units, and the symbols in the table show which units scored significantly higher or lower than each U.S. group—total, non-Hispanic white, Asian, Hispanic, and non-Hispanic Black.[8] Red-filled areas denote nations that scored higher than the U.S. population indicated at the top of the column. Green-filled areas denote nations that scored lower than the U.S. population indicated at the top of the column.

As shown in Table 1, U.S. Asians ranked fourth in the world, and U.S. whites ranked sixth. Both scored higher in reading literacy than the OECD average. Only Singapore ranked higher than U.S. non-Hispanic whites, and no nation ranked statistically higher than U.S. Asians. The nine-point difference between Singapore's top-ranked score and that of U.S. non-Hispanic whites was less than 0.10 standard deviations. Fifty-four populations averaged statistically lower than U.S. Asians. Among the 68 populations that scored below U.S. non-Hispanic whites, all except Ireland, Estonia, and the Republic of Korea scored statistically lower.

The situation is entirely different for U.S. Hispanics and U.S. non-Hispanic Blacks, both of which scored significantly lower than the OECD average. U.S. Hispanics ranked 40th in reading literacy. Thirty populations scored significantly higher, and 30 scored significantly lower. With a score of 443, U.S. non-Hispanic Blacks ranked 49th. That score was 0.87 standard deviations below that of U.S. non-Hispanic whites. Forty-five populations scored significantly higher, and only 21 scored significantly lower in reading literacy than U.S. non-Hispanic Blacks. The average score of U.S. Hispanics was 0.51 standard deviations lower than that of U.S. non-Hispanic whites. Simply put, U.S. non-Hispanic whites and U.S. Asians are world-class readers, while U.S. Hispanics and U.S. non-Hispanic Blacks lag much of the world.

Science

The story is much the same in PISA 2015 rankings on science literacy (Table 2). In science as in reading, the average achievement of U.S.

7 Tables 1, 2, and 5 are modifications of Tables R1, S1, and M1 in the NCES spreadsheets, based on information on scores by race-ethnicity in Tables R5, S10, and M5, respectively. I use the term *population* to refer to participating units in PISA, for some are nations, city-states, or provinces, and some are U.S. population groups.

8 Statistical significance is set here at the 5 percent level—the same as that used to report on national differences by the NCES. However, given the relatively large size of the national samples in PISA, it might have been preferable to set a higher standard for statistical significance, e.g., 1 percent, in which case only larger differences among population groups would be treated as reliable.

non-Hispanic whites and U.S. Asians is world-class, while the lower—but by no means bottom-ranking—scores of U.S. Hispanics and U.S. non-Hispanic Blacks substantially reduce the overall standing of U.S. students. Among the 74 populations, U.S. non-Hispanic whites averaged 531, which is 38 points higher than the OECD average, and were outranked only by Singapore, Japan, Estonia, and Taiwan.[9] U.S. Asians scored only six points lower and were ranked ninth. U.S. non-Hispanic whites scored significantly higher than 63 populations, and U.S. Asians scored significantly higher than 47 populations.

U.S. Hispanics ranked 42nd and scored 470, which is 23 points lower than the OECD average, while U.S. non-Hispanic Blacks ranked 52nd with an average score of 433, which is 60 points lower than the OECD average and a full standard deviation below the score of U.S. non-Hispanic whites. U.S. non-Hispanic Blacks scored significantly higher than only 17 populations. U.S. Hispanics scored significantly higher than 29 populations, but 0.64 standard deviations below U.S. non-Hispanic whites.

Another international comparative assessment, TIMSS, provides evidence that is generally consistent with PISA. In 2015, TIMSS tested fourth- and eighth-grade students. Thus, the study populations were defined by grade level, rather than age. Table 3 shows average TIMSS science scores and rankings of fourth-grade students in 51 populations.[10] Across grade levels and subjects, the standard deviation of TIMMS scores in the United States ranged from 80 to 85. U.S. Asians scored 598—higher than any other population group—and U.S. non-Hispanic whites scored 570, which ranked fourth among 51 populations. The U.S. average, 546, ranked 12th. However, U.S. Hispanics ranked 30th at 518, and U.S. non-Hispanic Blacks ranked 37th at 501, which is almost exactly the centerpoint (500) of the international scale.

In the eighth-grade TIMSS science assessment (Table 4), the U.S. total score of 530 ranked 13th among 41 populations. U.S. Asians' score of 573 trailed only Singapore, and U.S. non-Hispanic whites' score of 557 ranked fifth. However, the U.S. Hispanic score of 502 was close to the centerpoint of the TIMSS scale (500), and the U.S. non-Hispanic Black score of 469 ranked 28th and was significantly lower than the centerpoint.

---

9 For political reasons, PISA lists Taiwan as Taipei.

10 Tables 3, 4, 6, and 7 are modifications of Tables 23, 24, 1, and 2 in the NCES spreadsheet, alltables.xlsx (https://nces.ed.gov/timss/timss2015/), based on information on scores by race-ethnicity in Tables 41, 42, 19, and 20, respectively.

Rank	Education system	Average score	s.e.	Total	White	Asian	Hispanic	Black
	OECD average	493	0.5		▽	▽	△	△
1	*Singapore*	535	1.6	△	△		△	△
2	*Hong Kong (China)*	527	2.7	△			△	△
3	Canada	527	2.3	△			△	△
4	**U.S. Asian**	**527**	***13.3***	△			△	△
5	Finland	526	2.5	△			△	△
6	**U.S. White**	**526**	***3.3***	△			△	△
7	Ireland	521	2.5	△			△	△
8	Estonia	519	2.2	△			△	△
9	Korea, Republic of	517	3.5	△			△	△
10	Japan	516	3.2	△	▽		△	△
11	Norway	513	2.5	△	▽		△	△
12	New Zealand	509	2.4	△	▽		△	△
13	Germany	509	3.0	△	▽		△	△
14	*Macau (China)*	509	1.3	△	▽		△	△
15	Poland	506	2.5	△	▽		△	△
16	Slovenia	505	1.5	△	▽		△	△
17	Netherlands	503	2.4		▽		△	△
18	Australia	503	1.7		▽		△	△
19	Sweden	500	3.5		▽		△	△
20	Denmark	500	2.5		▽		△	△
21	France	499	2.5		▽	▽	△	△
22	Belgium	499	2.4		▽	▽	△	△
23	Portugal	498	2.7		▽	▽	△	△
24	United Kingdom	498	2.8		▽	▽	△	△
25	*Chinese Taipei*	497	2.5		▽	▽	△	△
26	**United States**	**497**	***3.4***		▽	▽	△	△
27	Spain	496	2.4		▽	▽	△	△
28	*Russian Federation*	495	3.1		▽	▽	△	△
29	*B-S-J-G (China)*	494	5.1		▽	▽	△	△
30	Switzerland	492	3.0		▽	▽	△	△
31	Latvia	488	1.8	▽	▽	▽		△
32	Czech Republic	487	2.6	▽	▽	▽		△
33	*Croatia*	487	2.7	▽	▽	▽		△
34	*Vietnam*	487	3.7	▽	▽	▽		△
35	Austria	485	2.8	▽	▽	▽		△
36	Italy	485	2.7	▽	▽	▽		△
37	Iceland	482	2.0	▽	▽	▽		△

TABLE 1. Average Scores of 15-Year-Old Students on the PISA Reading Literacy Scale, by Education System: 2015. (Continued on next page)
Highlighted in red: Average score is higher than U.S. total or subgroup average score at the 0.05 level of statistical significance.
Highlighted in green: Average score is lower than U.S. total or subgroup average score at the 0.05 level of statistical significance.

Mathematics

The evidence of selective U.S. excellence is mixed in the case of mathematics. It is not as clear as in reading and science that U.S. non-Hispanic whites and U.S. Asians are world-class in mathematics achievement. However, their average performance is still respectable and far above that of U.S. Hispanics and U.S. non-Hispanic Blacks. At age 15, the PISA data ranked U.S. non-Hispanic whites and U.S. Asians as 20th and 21st among 74 populations with scores of 499 and 498 (Table 5). Even U.S. non-Hispanic whites averaged more than three-quarters of a standard deviation below world-leading Singapore. The U.S. total ranking is 42nd, and the score of 470 is significantly lower than the

Rank	Education system	Average score	s.e.	Total	White	Asian	Hispanic	Black
38	Luxembourg	481	1.4	▽	▽	▽		▲
39	Israel	479	3.8	▽	▽	▽		▲
40	**U.S. Hispanic**	**478**	**5.7**	▽	▽	▽		▲
41	*Buenos Aires (Argentina)*	475	7.2	▽	▽	▽		▲
42	*Lithuania*	472	2.7	▽	▽	▽		▲
43	Hungary	470	2.7	▽	▽	▽		▲
44	Greece	467	4.3	▽	▽	▽		▲
45	Chile	459	2.6	▽	▽	▽	▽	▲
46	Slovak Republic	453	2.8	▽	▽	▽	▽	▲
47	*Malta*	447	1.8	▽	▽	▽	▽	▲
48	*Cyprus*	443	1.7	▽	▽	▽	▽	▲
49	**U.S. Black**	**443**	**5.4**	▽	▽	▽	▽	▲
50	*Uruguay*	437	2.5	▽	▽	▽	▽	▽
51	*Romania*	434	4.1	▽	▽	▽	▽	▽
52	*United Arab Emirates*	434	2.9	▽	▽	▽	▽	▽
53	*Bulgaria*	432	5.0	▽	▽	▽	▽	▽
54	Turkey	428	4.0	▽	▽	▽	▽	▽
55	*Costa Rica*	427	2.6	▽	▽	▽	▽	▽
56	*Trinidad and Tobago*	427	1.5	▽	▽	▽	▽	▽
57	*Montenegro, Republic of*	427	1.6	▽	▽	▽	▽	▽
58	*Colombia*	425	2.9	▽	▽	▽	▽	▽
59	Mexico	423	2.6	▽	▽	▽	▽	▽
60	*Moldova, Republic of*	416	2.5	▽	▽	▽	▽	▽
61	*Thailand*	409	3.3	▽	▽	▽	▽	▽
62	*Jordan*	408	2.9	▽	▽	▽	▽	▽
63	*Brazil*	407	2.8	▽	▽	▽	▽	▽
64	*Albania*	405	4.1	▽	▽	▽	▽	▽
65	*Qatar*	402	1.0	▽	▽	▽	▽	▽
66	*Georgia*	401	3.0	▽	▽	▽	▽	▽
67	*Peru*	398	2.9	▽	▽	▽	▽	▽
68	*Indonesia*	397	2.9	▽	▽	▽	▽	▽
69	*Tunisia*	361	3.1	▽	▽	▽	▽	▽
70	*Dominican Republic*	358	3.1	▽	▽	▽	▽	▽
71	*Macedonia, Republic of*	352	1.4	▽	▽	▽	▽	▽
72	*Algeria*	350	3.0	▽	▽	▽	▽	▽
73	*Kosovo*	347	1.6	▽	▽	▽	▽	▽
74	*Lebanon*	347	4.4	▽	▽	▽	▽	▽

TABLE 1. (Continued)

Note: Education systems are ordered by 2015 average score. The OECD average is the average of the national averages of the OECD member countries, with each country weighted equally. Scores are reported on a scale from 0 to 1,000. All average scores reported as higher or lower than the U.S. average score are different at the 0.05 level of statistical significance. Italics indicate non-OECD countries and education systems. B-S-J-G (China) refers to the four PISA-participating China provinces: Beijing, Shanghai, Jiangsu, and Guangdong. Results for Massachusetts and North Carolina are for public school students only. Although Argentina, Malaysia, and Kazakhstan participated in PISA 2015, technical problems with their samples prevent results from being discussed in this report. This table corresponds to Table 2 in *Performance of U.S. 15-Year-Old Students in Science, Mathematics, and Reading Literacy in an International Context* (NCES 2017-048).

Source: OECD, PISA 2015.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
	OECD average	493	0.4		▼	▼	▲	▲
1	*Singapore*	556	1.2	▲	▲	▲	▲	▲
2	Japan	538	3.0	▲			▲	▲
3	Estonia	534	2.1	▲			▲	▲
4	*Taiwan*	532	2.7	▲			▲	▲
5	**U.S. White**	**531**	**2.8**	▲			▲	▲
6	Finland	531	2.4	▲			▲	▲
7	*Macau (China)*	529	1.1	▲			▲	▲
8	Canada	528	2.1	▲			▲	▲
9	**U.S. Asian**	**525**	**12.0**	▲			▲	▲
10	*Vietnam*	525	3.9	▲			▲	▲
11	*Hong Kong (China)*	523	2.5	▲	▼		▲	▲
12	*B-S-J-G (China)*	518	4.6	▲	▼		▲	▲
13	Korea, Republic of	516	3.1	▲	▼		▲	▲
14	New Zealand	513	2.4	▲	▼		▲	▲
15	Slovenia	513	1.3	▲	▼		▲	▲
16	Australia	510	1.5	▲	▼		▲	▲
17	United Kingdom	509	2.6	▲	▼		▲	▲
18	Germany	509	2.7	▲	▼		▲	▲
19	Netherlands	509	2.3	▲	▼		▲	▲
20	Switzerland	506	2.9	▲	▼		▲	▲
21	Ireland	503	2.4		▼		▲	▲
22	Belgium	502	2.3		▼		▲	▲
23	Denmark	502	2.4		▼		▲	▲
24	Poland	501	2.5		▼		▲	▲
25	Portugal	501	2.4		▼		▲	▲
26	Norway	498	2.3		▼	▼	▲	▲
27	**United States**	**496**	**3.2**		▼	▼	▲	▲
28	Austria	495	2.4		▼	▼	▲	▲
29	France	495	2.1		▼	▼	▲	▲
30	Sweden	493	3.6		▼	▼	▲	▲
31	Czech Republic	493	2.3		▼	▼	▲	▲
32	Spain	493	2.1		▼	▼	▲	▲
33	Latvia	490	1.6		▼	▼	▲	▲
34	*Russian Federation*	487	2.9	▼	▼	▼	▲	▲
35	Luxembourg	483	1.1	▼	▼	▼	▲	▲
36	Italy	481	2.5	▼	▼	▼	▲	▲
37	Hungary	477	2.4	▼	▼	▼		▲

TABLE 2. Average Scores of 15-Year-Old Students on the PISA Science Literacy Scale, by Education System: 2015. (Continued on next page)
Highlighted in red: Average score is higher than U.S. total or subgroup average score at the 0.05 level of statistical significance.
Highlighted in green: Average score is lower than U.S. total or subgroup average score at the 0.05 level of statistical significance.

OECD average of 490. In mathematics, the top seven spots are occupied by Asian nations or city-states: Singapore, Hong Kong, Macau, Taiwan, Japan, four Chinese provinces, and the Republic of Korea.[11] Again, U.S. Hispanic and U.S. Black averages were well down the list: U.S. Hispanics scored 446 and ranked 46th, and U.S. non-Hispanic Blacks scored 419 and ranked 54th. As in science, U.S. non-Hispanic

11 For political reasons, PISA lists Taiwan as Taipei.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
38	*Lithuania*	475	*2.7*	▼	▼	▼		▲
39	*Croatia*	475	*2.5*	▼	▼	▼		▲
40	*Buenos Aires (Argentina)*	475	*6.3*	▼	▼	▼		▲
41	Iceland	473	*1.7*	▼	▼	▼		▲
42	**U.S. Hispanic**	**470**	**4.8**	▼	▼	▼		▲
43	Israel	467	*3.4*	▼	▼	▼		▲
44	*Malta*	465	*1.6*	▼	▼	▼		▲
45	Slovak Republic	461	*2.6*	▼	▼	▼		▲
46	Greece	455	*3.9*	▼	▼	▼	▼	▲
47	Chile	447	*2.4*	▼	▼	▼	▼	▲
48	*Bulgaria*	446	*4.4*	▼	▼	▼	▼	
49	*United Arab Emirates*	437	*2.4*	▼	▼	▼	▼	
50	*Uruguay*	435	*2.2*	▼	▼	▼	▼	
51	*Romania*	435	*3.2*	▼	▼	▼	▼	
52	**U.S. Black**	**433**	**4.9**	▼	▼	▼	▼	
53	*Cyprus*	433	*1.4*	▼	▼	▼	▼	
54	*Moldova, Republic of*	428	*2.0*	▼	▼	▼	▼	
55	*Albania*	427	*3.3*	▼	▼	▼	▼	
56	Turkey	425	*3.9*	▼	▼	▼	▼	
57	*Trinidad and Tobago*	425	*1.4*	▼	▼	▼	▼	
58	*Thailand*	421	*2.8*	▼	▼	▼	▼	▼
59	*Costa Rica*	420	*2.1*	▼	▼	▼	▼	▼
60	*Qatar*	418	*1.0*	▼	▼	▼	▼	▼
61	*Colombia*	416	*2.4*	▼	▼	▼	▼	▼
62	Mexico	416	*2.1*	▼	▼	▼	▼	▼
63	*Montenegro, Republic of*	411	*1.0*	▼	▼	▼	▼	▼
64	*Georgia*	411	*2.4*	▼	▼	▼	▼	▼
65	*Jordan*	409	*2.7*	▼	▼	▼	▼	▼
66	*Indonesia*	403	*2.6*	▼	▼	▼	▼	▼
67	*Brazil*	401	*2.3*	▼	▼	▼	▼	▼
68	*Peru*	397	*2.4*	▼	▼	▼	▼	▼
69	*Lebanon*	386	*3.4*	▼	▼	▼	▼	▼
70	*Tunisia*	386	*2.1*	▼	▼	▼	▼	▼
71	*Macedonia, Republic of*	384	*1.2*	▼	▼	▼	▼	▼
72	*Kosovo*	378	*1.7*	▼	▼	▼	▼	▼
73	*Algeria*	376	*2.6*	▼	▼	▼	▼	▼
74	Dominican Republic	332	*2.6*	▼	▼	▼	▼	▼

TABLE 2. (Continued)

Note: Education systems are ordered by 2015 average score. The OECD average is the average of the national averages of the OECD member countries, with each country weighted equally. Scores are reported on a scale from 0 to 1,000. All average scores reported as higher or lower than the U.S. average score are different at the 0.05 level of statistical significance. Italics indicate non-OECD countries and education systems. B-S-J-G (China) refers to the four PISA-participating China provinces: Beijing, Shanghai, Jiangsu, and Guangdong. Results for Massachusetts and North Carolina are for public school students only. Although Argentina, Malaysia, and Kazakhstan participated in PISA 2015, technical problems with their samples prevent results from being discussed in this report. This table corresponds to Table 1 in *Performance of U.S. 15-Year-Old Students in Science, Mathematics, and Reading Literacy in an International Context* (NCES 2017-048).

Source: OECD, PISA 2015.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
	TIMSS scale centerpoint	500	0.0	▽	▽	▽	▽	
1	**U.S. Asian**	**598**	**8.1**		△		△	△
2	Singapore[1]	590	3.7	△	△		△	△
3	Korea, Republic of	589	2.0	△	△		△	△
4	**U.S. White**	**570**	**2.0**	△		▽	△	△
5	Japan	569	1.8	△		▽	△	△
6	Russian Federation	567	3.2	△		▽	△	△
7	Hong Kong -CHN[2]	557	2.9	△	▽	▽	△	△
8	Taiwan	555	1.8	△	▽	▽	△	△
9	Finland	554	2.3	△	▽	▽	△	△
10	Kazakhstan	550	4.4		▽	▽	△	△
11	Poland	547	2.4		▽	▽	△	△
12	**United States[1, 2]**	**546**	**2.2**		▽	▽	△	△
13	Slovenia	543	2.4		▽	▽	△	△
14	Hungary	542	3.3		▽	▽	△	△
15	Sweden[1]	540	3.6		▽	▽	△	△
16	Norway (5)[3]	538	2.6	▽	▽	▽	△	△
17	England-GBR	536	2.4	▽	▽	▽	△	△
18	Bulgaria	536	5.9		▽	▽	△	△
19	Czech Republic	534	2.4	▽	▽	▽	△	△
20	Croatia	533	2.1	▽	▽	▽	△	△
21	Ireland	529	2.4	▽	▽	▽	△	△
22	Germany	528	2.4	▽	▽	▽	△	△
23	Lithuania[1]	528	2.5	▽	▽	▽	△	△
24	Denmark[1, 2]	527	2.1	▽	▽	▽	△	△
25	Canada[1, 2, 4]	525	2.6	▽	▽	▽		△
26	Serbia[5]	525	3.7	▽	▽	▽		△

TABLE 3. Average Science Scores of Fourth-Grade Students, by Education System: TIMSS 2015. (Continued on next page)

Highlighted in red: Average score is higher than U.S. total or subgroup average score at the 0.05 level of statistical significance.

Highlighted in green: Average score is lower than U.S. total or subgroup average score at the 0.05 level of statistical significance.

[1] National Defined Population covers 90 to 95 percent of the National Target Population.

[2] Met guidelines for sample participation rates only after replacement schools were included.

[3] The number in parentheses indicates years of school, not grade in schooling.

[4] National Target Population does not include all of the International Target Population.

[5] National Defined Population covers less than 90 percent of the National Target Population (but at least 77 percent).

[6] Nearly satisfied guidelines for sample participation rates after replacement schools were included.

[7] Reservations about reliability because the percentage of students with achievement too low for estimation exceeds 15 percent but does not exceed 25 percent.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
27	Australia	524	2.9	▽	▽	▽		△
28	Slovak Republic	520	2.6	▽	▽	▽		△
29	Northern Ireland-GBR[6]	520	2.2	▽	▽	▽		△
30	**U.S. Hispanic**	**518**	**3.0**	▽	▽	▽		△
31	Spain[1]	518	2.6	▽	▽	▽		△
32	Netherlands[2]	517	2.7	▽	▽	▽		△
33	Italy[1]	516	2.6	▽	▽	▽		△
34	Belgium (Flemish)-BEL[2]	512	2.3	▽	▽	▽		△
35	Portugal[1]	508	2.2	▽	▽	▽	▽	
36	New Zealand	506	2.7	▽	▽	▽	▽	
37	**U.S. Black**	**501**	**3.5**	▽	▽	▽	▽	
38	France	487	2.7	▽	▽	▽	▽	▽
39	Turkey	483	3.3	▽	▽	▽	▽	▽
40	Cyprus	481	2.6	▽	▽	▽	▽	▽
41	Chile	478	2.7	▽	▽	▽	▽	▽
42	Bahrain[1]	459	2.6	▽	▽	▽	▽	▽
43	Georgia[4]	451	3.7	▽	▽	▽	▽	▽
44	United Arab Emirates	451	2.8	▽	▽	▽	▽	▽
45	Qatar	436	4.1	▽	▽	▽	▽	▽
46	Oman	431	3.1	▽	▽	▽	▽	▽
47	Iran, Islamic Republic of	421	4.0	▽	▽	▽	▽	▽
48	Indonesia	397	4.8	▽	▽	▽	▽	▽
49	Saudi Arabia	390	4.9	▽	▽	▽	▽	▽
50	Morocco[7]	352	4.7	▽	▽	▽	▽	▽
51	Kuwait[7]	337	6.2	▽	▽	▽	▽	▽

TABLE 3. (Continued)

Note: Education systems are ordered by average score. Education systems that are not countries are designated by the appended three-letter international abbreviation for their country. Participants that did not administer TIMSS at the target grade are not shown; see the international report for their results. U.S. state data are based on public school students only. The TIMSS scale centerpoint is set at 500 points and represents the mean of the overall achievement distribution in 1995. The TIMSS scale is the same in each administration; thus, a value of 500 in 2015 equals 500 in 1995. Standard error is abbreviated as s.e. For TIMSS 2015, Norway revised its assessed population to students in their fifth and ninth years of schooling to obtain better comparisons with Sweden and Finland. However, in previous TIMSS cycles Norway assessed students in their fourth and eighth years of schooling, which were defined as fourth and eighth grades but have been redefined as third and seventh grades because year 1 in Norway is now considered the equivalent of a year of kindergarten. To maintain trend with previous TIMSS cycles, in 2015 Norway also collected data from students in their fourth and eighth years of schooling, which is used in trend tables. Jordan did not participate in the science assessment of the fourth grade.

Source: International Association for the Evaluation of Educational Achievement (IEA), TIMSS 2015.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
	TIMSS scale centerpoint	500	0.0	▽	▽	▽		△
1	Singapore[1]	597	3.2	△	△	△	△	△
2	**U.S. Asian**	**573**	**7.4**	△	△		△	△
3	Japan	571	1.8	△	△		△	△
4	Taiwan	569	2.1	△	△		△	△
5	**U.S. White**	**557**	**2.5**	△		▽	△	△
6	Korea, Rep. of	556	2.3	△		▽	△	△
7	Slovenia	551	2.4	△		▽	△	△
8	Hong Kong-CHN	546	3.9	△	▽	▽	△	△
9	Russian Federation	544	4.2	△	▽	▽	△	△
10	England-GBR	537	3.8		▽	▽	△	△
11	Kazakhstan	533	4.4		▽	▽	△	△
12	Ireland	530	2.8		▽	▽	△	△
13	**United States[2]**	**530**	**2.8**		▽	▽	△	△
14	Hungary	527	3.4		▽	▽	△	△
15	Canada[2, 3]	526	2.1		▽	▽	△	△
16	Sweden	522	3.5		▽	▽	△	△
17	Lithuania[1]	519	2.8	▽	▽	▽	△	△
18	New Zealand[2]	513	3.1	▽	▽	▽	△	△
19	Australia	512	2.7	▽	▽	▽	△	△
20	Norway (9)[4]	509	2.8	▽	▽	▽		△
21	Israel[5]	507	3.9	▽	▽	▽		△

TABLE 4. Average Science Scores of Eighth-Grade Students, by Education System: TIMSS 2015. (Continued on next page)
Highlighted in red: Average score is higher than U.S. total or subgroup average score at the 0.05 level of statistical significance.
Highlighted in green: Average score is lower than U.S. total or subgroup average score at the 0.05 level of statistical significance.
[1] National Defined Population covers 90 to 95 percent of the National Target Population.
[2] Met guidelines for sample participation rates only after replacement schools were included.
[3] National Target Population does not include all of the International Target Population.
[4] The number in parentheses indicates years of school, not grade in schooling.
[5] National Defined Population covers less than 90 percent of the National Target Population (but at least 77 percent).

Blacks and U.S. Hispanics, respectively, averaged about one standard deviation and two-thirds of a standard deviation below U.S. non-Hispanic whites.

At the same time, the 2015 TIMSS data in Tables 6 and 7 offer a more positive picture of the math achievement of U.S. Asians and U.S. non-Hispanic whites. In the fourth grade, the U.S. total average ranked 16th among 52 populations; the score of 539 is well above the TIMSS scale centerpoint of 500. U.S. Asians averaged 605 on the TIMSS scale

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
22	**U.S. Hispanic**	**502**	**3.5**	▽	▽	▽		△
23	Italy[1]	499	2.4	▽	▽	▽		△
24	Turkey	493	4.0	▽	▽	▽		△
25	Malta	481	1.6	▽	▽	▽	▽	△
26	United Arab Emirates	477	2.3	▽	▽	▽	▽	
27	Malaysia	471	4.1	▽	▽	▽	▽	
28	**U.S. Black**	**469**	**4.7**	▽	▽	▽	▽	
29	Bahrain	466	2.2	▽	▽	▽	▽	
30	Qatar	457	3.0	▽	▽	▽	▽	▽
31	Iran, Islamic Rep. of	456	4.0	▽	▽	▽	▽	▽
32	Thailand	456	4.2	▽	▽	▽	▽	▽
33	Oman	455	2.7	▽	▽	▽	▽	▽
34	Chile	454	3.1	▽	▽	▽	▽	▽
35	Georgia[1, 3]	443	3.1	▽	▽	▽	▽	▽
36	Jordan	426	3.3	▽	▽	▽	▽	▽
37	Kuwait	411	5.2	▽	▽	▽	▽	▽
38	Lebanon	398	5.3	▽	▽	▽	▽	▽
39	Saudi Arabia	396	4.5	▽	▽	▽	▽	▽
40	Morocco	393	2.5	▽	▽	▽	▽	▽
41	Egypt	371	4.3	▽	▽	▽	▽	▽

TABLE 4. (Continued)
Note: Education systems are ordered by average score. Education systems that are not countries are designated by the appended three-letter international abbreviation for their country. Participants that did not administer TIMSS at the target grade are not shown; see the international report for their results. U.S. state data are based on public school students only. The TIMSS scale centerpoint is set at 500 points and represents the mean of the overall achievement distribution in 1995. The TIMSS scale is the same in each administration; thus, a value of 500 in 2015 equals 500 in 1995. Standard error is abbreviated as s.e. For TIMSS 2015, Norway revised its assessed population to students in their fifth and ninth years of schooling to obtain better comparisons with Sweden and Finland. However, in previous TIMSS cycles Norway assessed students in their fourth and eighth years of schooling, which were defined as fourth and eighth grades but have been redefined as third and seventh grades because year 1 in Norway is now considered the equivalent of a year of kindergarten. To maintain trend with previous TIMSS cycles, in 2015 Norway also collected data from students in their fourth and eighth years of schooling, which is used in trend tables.
Source: IEA, TIMSS 2015.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
	OECD average	490	0.4	▲	▼		▲	▲
1	*Singapore*	564	1.5	▲	▲	▲		▲
2	*Hong Kong (China)*	548	3.0	▲	▲	▲	▲	▲
3	*Macau (China)*	544	1.1	▲	▲	▲	▲	▲
4	*Taiwan*	542	3.0	▲	▲	▲	▲	▲
5	Japan	532	3.0	▲	▲	▲	▲	▲
6	*B-S-J-G (China)*	531	4.9	▲	▲	▲	▲	▲
7	Korea, Republic of	524	3.7	▲	▲	▲	▲	▲
8	Switzerland	521	2.9	▲	▲	▲	▲	▲
9	Estonia	520	2.0	▲	▲	▲	▲	▲
10	Canada	516	2.3	▲	▲		▲	▲
11	Netherlands	512	2.2	▲	▲		▲	▲
12	Denmark	511	2.2	▲	▲		▲	▲
13	Finland	511	2.3	▲	▲		▲	▲
14	Slovenia	510	1.3	▲	▲		▲	▲
15	Belgium	507	2.4	▲	▲		▲	▲
16	Germany	506	2.9	▲			▲	▲
17	Poland	504	2.4	▲			▲	▲
18	Ireland	504	2.1	▲			▲	▲
19	Norway	502	2.2	▲			▲	▲
20	**U.S. White**	499	2.8	▲			▲	▲
21	**U.S. Asian**	498	10.1	▲			▲	▲
22	Austria	497	2.9	▲			▲	▲
23	New Zealand	495	2.3	▲			▲	▲
24	*Vietnam*	495	4.5	▲			▲	▲
25	*Russian Federation*	494	3.1	▲			▲	▲
26	Sweden	494	3.2	▲			▲	▲
27	Australia	494	1.6	▲			▲	▲
28	France	493	2.1	▲			▲	▲
29	United Kingdom	492	2.5	▲			▲	▲
30	Czech Republic	492	2.4	▲			▲	▲
31	Portugal	492	2.5	▲			▲	▲
32	Italy	490	2.8	▲	▼		▲	▲
33	Iceland	488	2.0	▲	▼		▲	▲
34	Spain	486	2.2	▲	▼		▲	▲
35	Luxembourg	486	1.3	▲	▼		▲	▲
36	Latvia	482	1.9	▲	▼		▲	▲
37	*Malta*	479	1.7	▲	▼		▲	▲

TABLE 5. Average Scores of 15-Year-Old Students on the PISA Mathematics Literacy Scale, by Education System: 2015. (Continued on next page)
Highlighted in red: Average score is higher than U.S. total or subgroup average score at the 0.05 level of statistical significance.
Highlighted in green: Average score is lower than U.S. total or subgroup average score at the 0.05 level of statistical significance.

and ranked fourth, behind only Singapore, Hong Kong, and the Republic of Korea. U.S. non-Hispanic whites scored 559, significantly lower than U.S. Asians, but still ranked ninth overall. U.S. Hispanics scored 515, significantly above the TIMSS scale centerpoint, but ranked only 31st, and U.S. non-Hispanic Blacks scored 495, placing them 37th in the rank order.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
38	*Lithuania*	478	2.3	▲	▼		▲	▲
39	Hungary	477	2.5		▼	▼	▲	▲
40	Slovak Republic	475	2.7		▼	▼	▲	▲
41	Israel	470	3.6		▼	▼	▲	▲
42	**United States**	**470**	**3.2**		▼	▼	▲	▲
43	*Croatia*	464	2.8		▼	▼	▲	▲
44	*Buenos Aires (Argentina)*	456	6.9		▼	▼		▲
45	Greece	454	3.8	▼	▼	▼		▲
46	**U.S. Hispanic**	**446**	**5.2**	▼	▼	▼		▲
47	*Romania*	444	3.8	▼	▼	▼		▲
48	*Bulgaria*	441	4.0	▼	▼	▼		▲
49	*Cyprus*	437	1.7	▼	▼	▼		▲
50	*United Arab Emirates*	427	2.4	▼	▼	▼	▼	
51	Chile	423	2.5	▼	▼	▼	▼	
52	Turkey	420	4.1	▼	▼	▼	▼	
53	*Moldova, Republic of*	420	2.5	▼	▼	▼	▼	
54	**U.S. Black**	**419**	**4.7**	▼	▼	▼	▼	
55	*Uruguay*	418	2.5	▼	▼	▼	▼	
56	*Montenegro, Republic of*	418	1.5	▼	▼	▼	▼	
57	*Trinidad and Tobago*	417	1.4	▼	▼	▼	▼	
58	*Thailand*	415	3.0	▼	▼	▼	▼	
59	*Albania*	413	3.4	▼	▼	▼	▼	
60	Mexico	408	2.2	▼	▼	▼	▼	▼
61	*Georgia*	404	2.8	▼	▼	▼	▼	▼
62	*Qatar*	402	1.3	▼	▼	▼	▼	▼
63	*Costa Rica*	400	2.5	▼	▼	▼	▼	▼
64	*Lebanon*	396	3.7	▼	▼	▼	▼	▼
65	*Colombia*	390	2.3	▼	▼	▼	▼	▼
66	*Peru*	387	2.7	▼	▼	▼	▼	▼
67	*Indonesia*	386	3.1	▼	▼	▼	▼	▼
68	*Jordan*	380	2.7	▼	▼	▼	▼	▼
69	*Brazil*	377	2.9	▼	▼	▼	▼	▼
70	*Macedonia, Republic of*	371	1.3	▼	▼	▼	▼	▼
71	*Tunisia*	367	3.0	▼	▼	▼	▼	▼
72	*Kosovo*	362	1.6	▼	▼	▼	▼	▼
73	*Algeria*	360	3.0	▼	▼	▼	▼	▼
74	*Dominican Republic*	328	2.7	▼	▼	▼	▼	▼

TABLE 5. (Continued)
Note: Education systems are ordered by 2015 average score. The OECD average is the average of the national averages of the OECD member countries, with each country weighted equally. Scores are reported on a scale from 0 to 1,000. Standard error is abbreviated as s.e. Italics indicate non-OECD countries and education systems. B-S-J-G (China) refers to the four PISA participating China provinces: Beijing, Shanghai, Jiangsu, and Guangdong. Results for Massachusetts and North Carolina are for public school students only. Although Argentina, Malaysia, and Kazakhstan participated in PISA 2015, technical problems with their samples prevent results from being discussed in this report. This table corresponds to Table 3 in *Performance of U.S. 15-Year-Old Students in Science, Mathematics, and Reading Literacy in an International Context* (NCES 2017-048).
Source: OECD, PISA 2015.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
	TIMSS scale centerpoint	500	0.0	▽	▽	▽	▽	
1	Singapore[1]	618	3.8	△	△		△	△
2	Hong Kong-CHN[2]	615	2.9	△	△		△	△
3	Korea, Rep. of	608	2.2	△	△		△	△
4	**U.S. Asian**	**605**	**8.6**	△	△		△	△
5	Taiwan	597	1.9	△	△		△	△
6	Japan	593	2.0	△	△		△	△
7	Northern Ireland-GBR[3]	570	2.9	△	△	▽	△	△
8	Russian Federation	564	3.4	△		▽	△	△
9	**U.S. White**	**559**	**2.1**	△		▽	△	△
10	Norway (5)[4]	549	2.5	△	▽	▽	△	△
11	Ireland	547	2.1	△	▽	▽	△	△
12	England-GBR	546	2.8		▽	▽	△	△
13	Belgium (Flemish)-BEL[2]	546	2.1	△	▽	▽	△	△
14	Kazakhstan	544	4.5		▽	▽	△	△
15	Portugal[1]	541	2.2		▽	▽	△	△
16	**United States[1, 2]**	**539**	**2.3**		▽	▽	△	△
17	Denmark[1, 2]	539	2.7		▽	▽	△	△
18	Lithuania[1]	535	2.5		▽	▽	△	△
19	Finland	535	2.0		▽	▽	△	△
20	Poland	535	2.1		▽	▽	△	△
21	Netherlands[2]	530	1.7	▽	▽	▽	△	△
22	Hungary	529	3.2	▽	▽	▽	△	△
23	Czech Republic	528	2.2	▽	▽	▽	△	△
24	Bulgaria	524	5.3	▽	▽	▽		△
25	Cyprus	523	2.7	▽	▽	▽		△
26	Germany	522	2.0	▽	▽	▽		△

TABLE 6. Average Mathematics Scores of Fourth-Grade Students, by Education System: TIMSS 2015. (Continued on next page)

Highlighted in red: Average score is higher than U.S. total or subgroup average score at the 0.05 level of statistical significance.

Highlighted in green: Average score is lower than U.S. total or subgroup average score at the 0.05 level of statistical significance.

[1] National Defined Population covers 90 to 95 percent of the National Target Population.

[2] Met guidelines for sample participation rates only after replacement schools were included.

[3] Nearly satisfied guidelines for sample participation rates after replacement schools were included.

[4] The number in parentheses indicates years of school, not grade in schooling.

[5] National Defined Population covers less than 90 percent of the National Target Population (but at least 77 percent).

[6] National Target Population does not include all of the International Target Population.

[7] Reservations about reliability because the percentage of students with achievement too low for estimation exceeds 15 percent but does not exceed 25 percent.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
27	Slovenia	520	1.9	▽	▽	▽		△
28	Sweden[1]	519	2.8	▽	▽	▽		△
29	Serbia[5]	518	3.5	▽	▽	▽		△
30	Australia	517	3.1	▽	▽	▽		△
31	**U.S. Hispanic**	**515**	**3.1**	▽	▽	▽		△
32	Canada[1, 2, 6]	511	2.3	▽	▽	▽		△
33	Italy[1]	507	2.6	▽	▽	▽		△
34	Spain[1]	505	2.5	▽	▽	▽	▽	△
35	Croatia	502	1.8	▽	▽	▽	▽	
36	Slovak Republic	498	2.5	▽	▽	▽	▽	
37	**U.S. Black**	**495**	**3.6**	▽	▽	▽	▽	
38	New Zealand	491	2.3	▽	▽	▽	▽	
39	France	488	2.9	▽	▽	▽	▽	
40	Turkey	483	3.1	▽	▽	▽	▽	▽
41	Georgia[6]	463	3.6	▽	▽	▽	▽	▽
42	Chile	459	2.4	▽	▽	▽	▽	▽
43	United Arab Emirates	452	2.4	▽	▽	▽	▽	▽
44	Bahrain[1]	451	1.6	▽	▽	▽	▽	▽
45	Qatar	439	3.4	▽	▽	▽	▽	▽
46	Iran, Islamic Rep. of	431	3.2	▽	▽	▽	▽	▽
47	Oman	425	2.5	▽	▽	▽	▽	▽
48	Indonesia	397	3.7	▽	▽	▽	▽	▽
49	Jordan	388	3.1	▽	▽	▽	▽	▽
50	Saudi Arabia[7]	383	4.1	▽	▽	▽	▽	▽
51	Morocco	377	3.4	▽	▽	▽	▽	▽
52	Kuwait[7]	353	4.6	▽	▽	▽	▽	▽

TABLE 6. (Continued)
Note: Education systems are ordered by average score. Education systems that are not countries are designated by the appended three-letter international abbreviation for their country. Participants that did not administer TIMSS at the target grade are not shown; see the international report for their results. U.S. state data are based on public school students only. The TIMSS scale centerpoint is set at 500 points and represents the mean of the overall achievement distribution in 1995. The TIMSS scale is the same in each administration; thus, a value of 500 in 2015 equals 500 in 1995. Standard error is abbreviated as s.e. For TIMSS 2015, Norway revised its assessed population to students in their fifth and ninth years of schooling to obtain better comparisons with Sweden and Finland. However, in previous TIMSS cycles Norway assessed students in their fourth and eighth years of schooling, which were defined as fourth and eighth grades but have been redefined as third and seventh grades because year 1 in Norway is now considered the equivalent of a year of kindergarten. To maintain trend with previous TIMSS cycles, in 2015 Norway also collected data from students in their fourth and eighth years of schooling, which is used in trend tables.
Source: IEA, TIMSS 2015.

Rank	Education system	Average score	s.e.	US-T	US-W	US-A	US-H	US-B
	TIMSS scale centerpoint	500	0.0	▽	▽	▽	△	△
1	Singapore[1]	621	3.2	△	△	△	△	△
2	Korea, Rep. of	606	2.6	△	△	△	△	△
3	Taiwan	599	2.4	△	△		△	△
4	Hong Kong-CHN	594	4.6	△	△		△	△
5	Japan	586	2.3	△	△		△	△
6	**U.S. Asian**	**585**	8.5	△	△		△	△
7	**U.S. White**	**541**	3.0	△		▽	△	△
8	Russian Federation	538	4.7	△		▽	△	△
9	Kazakhstan	528	5.3		▽	▽	△	△
10	Canada[2, 3]	527	2.2	△	▽	▽	△	△
11	Ireland	523	2.7		▽	▽	△	△
12	**United States[3]**	**518**	3.1		▽	▽	△	△
13	England-GBR	518	4.2		▽	▽	△	△
14	Slovenia	516	2.1		▽	▽	△	△
15	Hungary	514	3.8		▽	▽	△	△
16	Norway (9)[4]	512	2.3		▽	▽	△	△
17	Lithuania[1]	511	2.8		▽	▽	△	△
18	Israel[5]	511	4.1		▽	▽	△	△
19	Australia	505	3.1	▽	▽	▽	△	△
20	Sweden	501	2.8	▽	▽	▽		△
21	Italy[1]	494	2.5	▽	▽	▽		△

TABLE 7. Average Mathematics Scores of Eighth-Grade Students, by Education System: TIMSS 2015. (Continued on next page)

Highlighted in red: Average score is higher than U.S. total or subgroup average score at the 0.05 level of statistical significance.

Highlighted in green: Average score is lower than U.S. total or subgroup average score at the 0.05 level of statistical significance.

[1] National Defined Population covers 90 to 95 percent of the National Target Population.

[2] National Target Population does not include all of the International Target Population.

[3] Met guidelines for sample participation rates only after replacement schools were included.

[4] The number in parentheses indicates years of school, not grade in schooling.

[5] National Defined Population covers less than 90 percent of the National Target Population (but at least 77 percent).

[6] Reservations about reliability because the percentage of students with achievement too low for estimation exceeds 15 percent but does not exceed 25 percent.

[7] Reservations about reliability because the percentage of students with achievement too low for estimation exceeds 25 percent.

Rank	Education system	Average score	s.e	US-T	US-W	US-A	US-H	US-B
22	Malta	494	1.0	▽	▽	▽		△
23	New Zealand[3]	493	3.4	▽	▽	▽		△
24	**U.S. Hispanic**	**492**	**3.8**	▽	▽	▽		△
25	Malaysia	465	3.6	▽	▽	▽	▽	
26	United Arab Emirates	465	2.0	▽	▽	▽	▽	
27	**U.S. Black**	**462**	**4.9**	▽	▽	▽	▽	
28	Turkey	458	4.7	▽	▽	▽	▽	
29	Bahrain	454	1.4	▽	▽	▽	▽	
30	Georgia[1, 2]	453	3.4	▽	▽	▽	▽	
31	Lebanon	442	3.6	▽	▽	▽	▽	▽
32	Qatar[6]	437	3.0	▽	▽	▽	▽	▽
33	Iran, Islamic Rep. of[6]	436	4.6	▽	▽	▽	▽	▽
34	Thailand	431	4.8	▽	▽	▽	▽	▽
35	Chile[6]	427	3.2	▽	▽	▽	▽	▽
36	Oman[6]	403	2.4	▽	▽	▽	▽	▽
37	Kuwait[6]	392	4.6	▽	▽	▽	▽	▽
38	Egypt[6]	392	4.1	▽	▽	▽	▽	▽
39	Jordan[7]	386	3.2	▽	▽	▽	▽	▽
40	Morocco[7]	384	2.3	▽	▽	▽	▽	▽
41	Saudi Arabia[7]	368	4.6	▽	▽	▽	▽	▽

TABLE 7. (Continued)
Note: Education systems are ordered by average score. Education systems that are not countries are designated by the appended three-letter international abbreviation for their country. Participants that did not administer TIMSS at the target grade are not shown; see the international report for their results. U.S. state data are based on public school students only. The TIMSS scale centerpoint is set at 500 points and represents the mean of the overall achievement distribution in 1995. The TIMSS scale is the same in each administration; thus, a value of 500 in 2015 equals 500 in 1995. Standard error is abbreviated as s.e. For TIMSS 2015, Norway revised its assessed population to students in their fifth and ninth years of schooling to obtain better comparisons with Sweden and Finland. However, in previous TIMSS cycles Norway assessed students in their fourth and eighth years of schooling, which were defined as fourth and eighth grades but have been redefined as third and seventh grades because year 1 in Norway is now considered the equivalent of a year of kindergarten. To maintain trend with previous TIMSS cycles, in 2015 Norway also collected data from students in their fourth and eighth years of schooling, which is used in trend tables.
Source: IEA, TIMSS 2015.

In the eighth grade, the U.S. total average ranked 12th among 41 populations, with the average score of 518, significantly higher than the TIMSS scale centerpoint of 500. U.S. Asians averaged 585 on the TIMSS scale and ranked sixth, only lower than five national or regional Asian populations. U.S. non-Hispanic whites averaged 541, significantly lower than U.S. Asians, but still ranked seventh overall. Again, U.S. Hispanic and U.S. Black averages were significantly lower. U.S. Hispanics scored 492, just below the TIMSS scale centerpoint, and ranked 24th, while U.S. non-Hispanic Blacks scored 462, placing them 27th in the rank order.

High and Low Scorers

The average (mean) scores reported in Tables 1 to 7 are a measure of what is typical in the distribution of achievement test scores in a nation, city-state, or other population group. Scores are also highly variable in every population. All populations include high and low scorers, regardless of their average score. Group differences in the relative frequency of high and low scores provide additional information. The NCES reports of 2015 PISA data include distributions of scores by race-ethnicity across pre-determined levels of proficiency in each assessment of literacy.[12] Each test item is assigned to a proficiency level, and each proficiency level is defined by a specific score range. To reach a given proficiency level, a student must supply the correct answer to the majority of items assigned to that level.

In reading literacy, PISA recognizes eight levels of proficiency (below 1b, 1b, 1a, 2, 3, 4, 5, and 6). Nineteen percent of all U.S. 15-year-old students scored at or below level 1a, which is close to the OECD figure of 20.1 percent. The upper limit of 1a is a score of 404, about the average in Albania, which ranks 64th internationally. In the top-ranking nation of Singapore, 11.1 percent of students scored at or below level 1a. Among U.S. non-Hispanic whites, 10.3 percent were at or below level 1a, along with 15 percent of U.S. Asians. However, 24.5 percent of U.S. Hispanics and 33.8 percent of U.S. non-Hispanic Blacks were at or below level 1a.

At the high end of the test score distribution, 30.1 percent of all U.S. students and 28.8 percent of all in the OECD scored at or above level 4. That proficiency level requires a score of 553, which lies above the average in any population group. In Singapore, 45.7 percent of

12 These are reported in Tables R2, R6, S11, S2b, S4, M2, and M6. Reported estimates are approximate because the source tables do not report estimated percentages in sparse cells, e.g., very low score ranges among U.S. non-Hispanic whites and U.S. Asians and very high score ranges among U.S. Hispanics and U.S. non-Hispanic Blacks.

students scored at or above level 4. That level was reached by 40.2 percent of U.S. non-Hispanic whites and 41.7 percent of U.S. Asians, compared with 23 percent of U.S. Hispanics and 10.5 percent of U.S. non-Hispanic Blacks.

In science literacy, PISA recognizes the same levels of proficiency as in reading literacy, but the score ranges of the levels are different. Of all U.S. 15-year-old students, 20.3 percent scored at or below level 1a, while the OECD figure is 21.2 percent. The upper limit of 1a is a score of 410, about the average in Jordan, which ranks 65th internationally. In Singapore, which is top-ranked in science as in reading, 9.6 percent of students scored at or below level 1a. Among U.S. non-Hispanic whites, 9.9 percent were at or below level 1a, along with 15.6 percent of U.S. Asians. However, 26.6 percent of U.S. Hispanics and 40.3 percent of U.S. non-Hispanic Blacks were at or below level 1a. At the high end of the test score distribution, 27.6 percent of all U.S. students and 26.8 percent of all in the OECD scored at or above level 4. That proficiency level requires a score of 559, which is just above the average in top-ranked Singapore, where 51.9 percent scored at or above level 4. Proficiency levels 4 or above were reached by 40 percent of U.S. non-Hispanic whites and 40.1 percent of U.S. Asians, compared with 16.7 percent of U.S. Hispanics and just 5.8 percent of U.S. non-Hispanic Blacks.

PISA 2015 defined only seven proficiency levels in mathematics literacy (below 1, 1, 2, 3, 4, 5, and 6). Level 1 has an upper score of 420, which is approximately the mean in Turkey and the Republic of Moldova, which ranked 52nd and 53rd internationally, just below the mean of U.S. non-Hispanic Blacks. Thus, unsurprisingly, 51.1 percent of U.S. non-Hispanic Blacks scored at or below level 1 in mathematics literacy, compared with 39.2 percent of U.S. Hispanics. In top-ranked Singapore, only 7.5 percent of students scored at or below level 1. Among U.S. non-Hispanic whites, 16.9 percent were at or below level 1a, along with 19 percent of U.S. Asians. At high levels of mathematics literacy, 20.6 percent of all U.S. students and 29.3 percent of all in the OECD scored at or above level 4. That proficiency level requires a score of 545, which is close to the average in third-ranked Macau, where 50.9 percent scored at or above level 4. In mathematics literacy, proficiency levels 4 or above were reached by 30.3 percent of U.S. non-Hispanic whites and 32.1 percent of U.S. Asians, compared with 11.8 percent of U.S. Hispanics and just 4.4 percent of U.S. non-Hispanic Blacks.

Again, while every population includes individuals at every level of academic proficiency, racial-ethnic differences are evident at both the high and low ends of the test-score distributions. While U.S. non-Hispanic

whites and U.S. Asians are international leaders—or at least competitive—in reading, science, and mathematics literacy, U.S. Hispanic and U.S. non-Hispanic Black populations lag far behind, and that accounts for the mediocre overall performance of U.S. students in international context. Racial-ethnic differences appear both in group averages and in the shares of exceptionally high- and low-scoring students. If policy makers and the American population truly want the United States to excel academically relative to other nations, their primary goal must be to improve the academic performance of the nation's largest minorities. That cannot be accomplished merely by reducing the share of low scorers among minorities, for those groups also have few high scorers. Real progress will necessitate shifting entire achievement score distributions in minority populations.

Social Origins and Academic Achievement

Given the racial-ethnic differentials in U.S. academic achievement, an immediate question is whether social and economic origins and schooling experiences explain them. In order to address this question, I analyzed data for individuals who participated in the U.S. PISA assessments in 2012 and 2015. By combining data for these two test administrations, it was possible to increase the statistical reliability of the analyses, which was especially important in the case of minority groups.

Table 8 presents initial and adjusted mean differences in each achievement assessment between U.S. non-Hispanic whites and the other five racial-ethnic groups identified in the U.S. PISA data, along with estimated standard errors (s.e.) of those differences. The initial differences are as observed, except the year of the PISA assessment is controlled statistically. The adjusted differences also control several other variables: socioeconomic status, gender, age, nativity, grade level, grade repetition, public/private school, and size of place. The socioeconomic status measure is a composite, developed by PISA, which includes "parents' education, parents' occupations, a number of home possessions that can be taken as proxies for material wealth, and the number of books and other educational resources available in the home" (Organisation for Economic Cooperation and Development 2016).

To be sure, the PISA data do not permit a complete accounting of the social and economic factors and school experiences that may account for group differences in average academic achievement. For example, one might wish to include information about school resources and family circumstances, such as housing tenure, number of siblings, and whether the family includes two parents. However, they do include several of the most important factors. Because other explanatory

	Reading				Science				Math			
	Initial		Adjusted		Initial		Adjusted		Initial		Adjusted	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
Black	-79.6	5.3	-64.4	5.1	-93.9	5.2	-77.7	5.1	-82.5	4.8	-67.7	4.6
Hispanic	-45.3	4.0	-21.8	4.3	-63.1	4.0	-37.6	4.3	-52.0	3.7	-26.9	3.9
Asian	16.4	8.4	13.4	8.0	7.0	8.3	6.3	8.1	24.0	7.7	22.1	7.3
Multiracial	-15.5	7.7	-10.2	7.1	-22.6	7.6	-16.2	7.1	-20.1	7.1	-14.4	6.5
Other	-83.9	13.9	-57.8	12.9	-82.6	13.7	-59.1	13.0	-73.6	12.6	-50.7	11.8

	Percentage Explained	Percentage Explained	Percentage Explained
Black	19.1	17.2	17.9
Hispanic	51.9	40.4	48.3
Asian	18.4	10.6	8.0
Multiracial	34.0	28.5	28.4
Other	31.2	28.5	31.1

TABLE 8. Observed and Adjusted Differences of Mean Test Scores of Race-Ethnic Groups from Non-Hispanic Whites: Pooled PISA 2012 and PISA 2015.

Note: All estimates are weighted. Initial differences adjust for survey year. Adjusted estimates control survey year, socioeconomic status (PISA scale), gender, age, nativity, grade level, grade repetition, public/private school, and size of place. Standard errors (s.e.) are rough estimates based on comparisons of regression estimates and published standard errors that account for weighting and sample design.

variables are undoubtedly correlated with the variables measured in PISA, the present analysis offers a reasonable approximation to what one would find in a more complete analysis.[13]

In summary, in no case does the array of explanatory variables available in the PISA data account for the differences in performance among the racial-ethnic groups. As measured, social background, school and community characteristics, and educational experiences account for less than 20 percent of the substantial differences in academic achievement between U.S. non-Hispanic Black and U.S. non-Hispanic white students. At the same time, those variables account for half of the differences between U.S. Hispanic and U.S. non-Hispanic white students in reading and mathematics and for 40 percent of the difference in science achievement. The findings for multiracial students and other students are midway between those for U.S. non-Hispanic Blacks and U.S. Hispanics; the explanatory variables account for about a third of the differential in each achievement domain. With or without the statistical controls, U.S. Asians outperform U.S. non-Hispanic whites—especially in mathematics—and the explanatory variables account for only a small share of the differences between them in academic achievement.

Trends in Achievement

Is it true, as some have suggested, that academic achievement in the United States has declined in recent years? In brief, while the trends in achievement in PISA are by no means uniform, there is no evidence of an overall systemic decline. Figures 1, 2, and 3 show the trends in available PISA data for each of the four largest racial-ethnic groups.[14] In reading, the performance of both minority groups, but not of U.S. non-Hispanic whites and U.S. Asians, is higher in 2009 and later years than in 2000 and 2003. In science, where the series covers only the years 2006 and beyond, the achievement of U.S. Hispanics and U.S. non-Hispanic Blacks rose between 2006 and later years. That of U.S. non-Hispanic whites was essentially stable, while that of U.S. Asians appears to have improved from 2006 to 2012 but declined slightly thereafter. In mathematics, there was some variation across years but no substantial change in any group between 2003 and 2015.

Similarly, the trend data from TIMSS show no evidence of score decline in any population group. Figures 4, 5, 6, and 7 show trend data

13 For example, more detailed analyses of group differences in academic achievement could be carried out using data from NCES longitudinal studies.

14 PISA reading data for 2006 are not comparable to those in other years and, thus, are omitted from Figure 1.

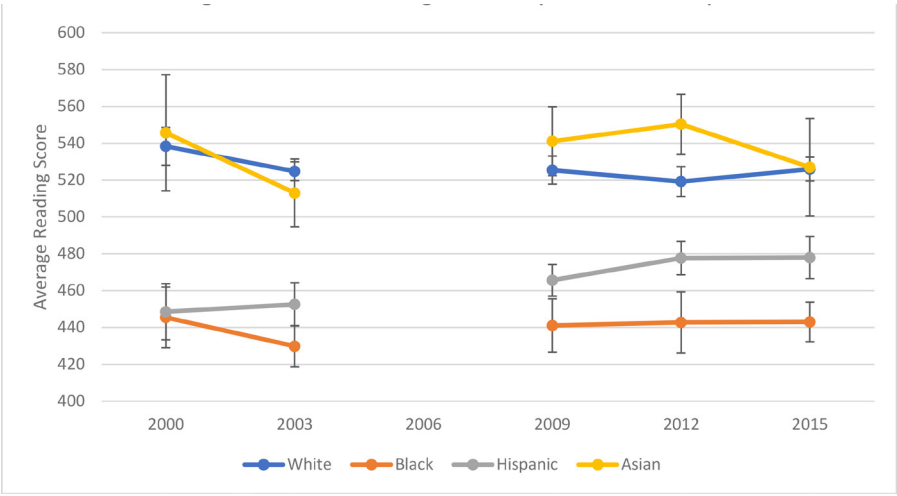

FIGURE 1. PISA Reading Trends by Race-Ethnicity.

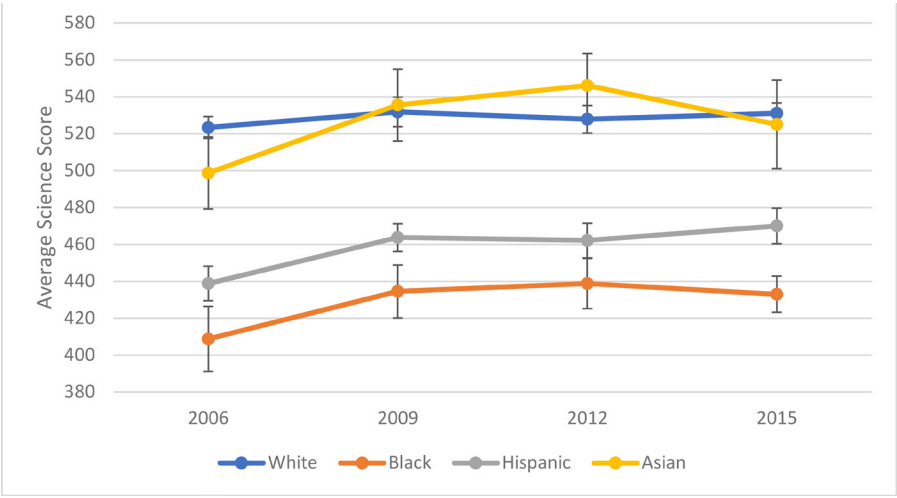

FIGURE 2. PISA Science Trends by Race-Ethnicity.

in average math and science scores from TIMSS assessments at grades 4 and 8 by race-ethnicity from 1995 to 2015. In mathematics, in both grades, there has been either growth or stability in average scores between each test administration. Average science scores in grade 4 rose among U.S. Asians and U.S. non-Hispanic Blacks between 1995 and later years, and the scores of U.S. Asians continued to rise through 2015. At the eighth-grade level, scores rose significantly among U.S. Asians, U.S. Hispanics, and U.S. non-Hispanic Blacks, but less so among U.S. non-Hispanic whites.

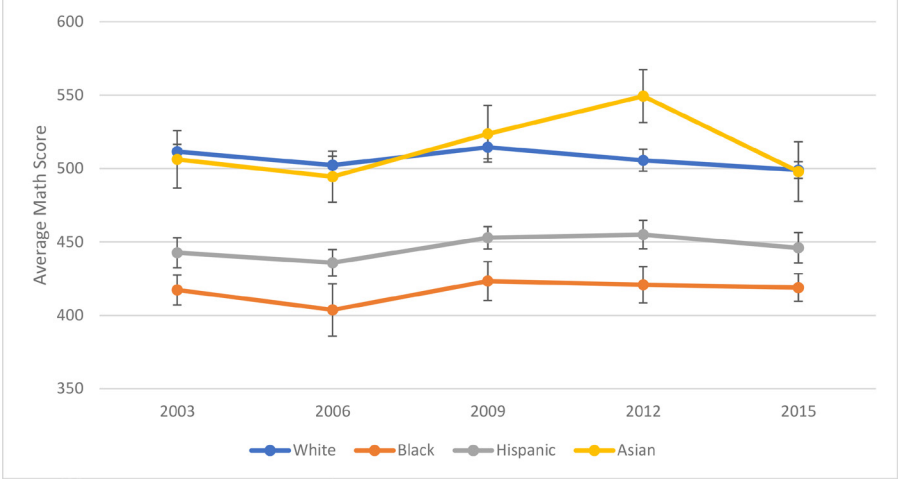

FIGURE 3. PISA Math Trends by Race-Ethnicity.

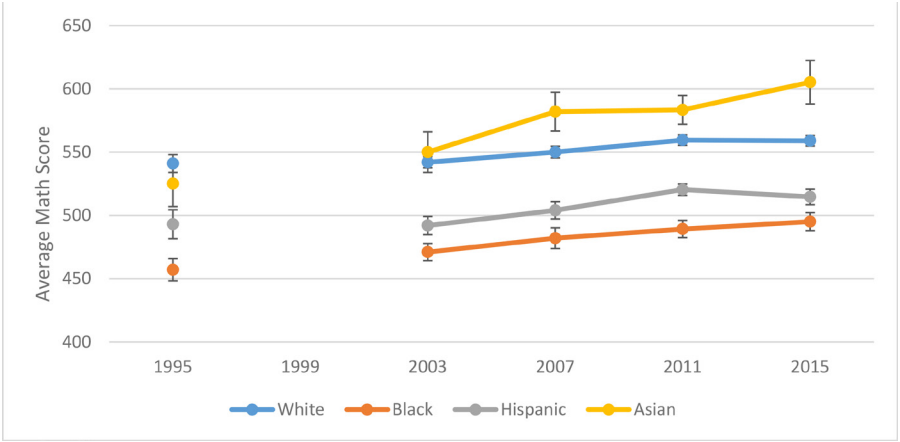

FIGURE 4. Average Math Scores in TIMSS: U.S. Fourth Graders by Race-Ethnicity, 1995–2015.

Figures 8, 9, 10, and 11 show trend data in average reading and mathematics scores at grades 4 and 8 by race and ethnicity from 2003 to 2017 in the NAEP, the premier, large-scale survey of academic achievement in the United States (https://nces.ed.gov/nationsreportcard/about/).[15] Again, there is no evidence of score decline. Unlike Figures 1–7, Figures 8–11 also show the average scores of all students

15 NAEP assessments are also administered in the 12th grade but are probably less valid for trend analysis than the assessments at the fourth- and eighth-grade levels. By the time they are in the 12th grade, a time-varying share of students have dropped out of school, and many students know that the NAEP assessments "do not count." NAEP also assesses science achievement, but it has done so only since 2009.

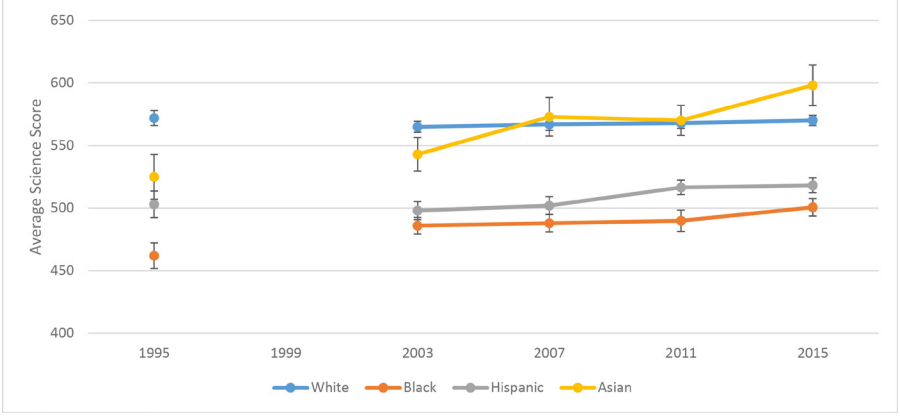

FIGURE 5. Average Science Scores in TIMSS: U.S. Fourth Graders by Race-Ethnicity, 1995–2015.

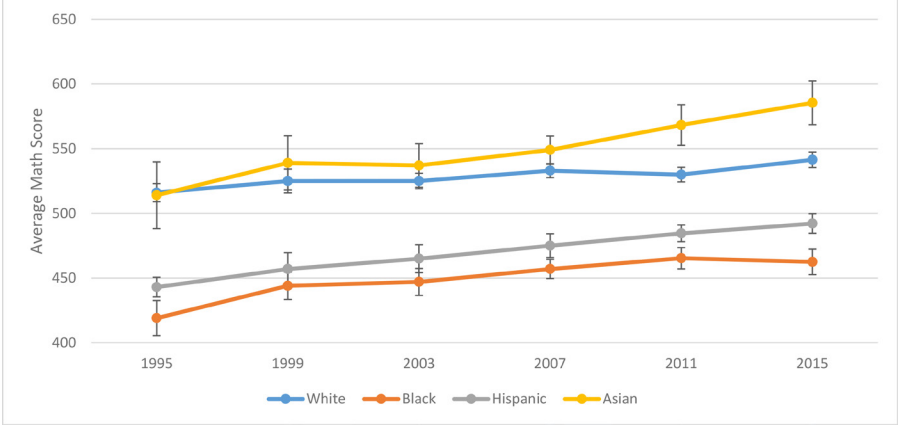

FIGURE 6. Average Math Scores in TIMSS: U.S. Eighth Graders by Race-Ethnicity, 1995–2015.

in each subject at each grade level, including smaller racial-ethnic groups. In reading at grade 4, there has been a very modest overall increase in average scores from 2003 to 2017 (Figure 8), while the scores of the three large minority groups—U.S. non-Hispanic Blacks, U.S. Hispanics, and U.S. Asians—have each increased substantially. The pattern is much the same for reading scores at the eighth-grade level (Figure 9). In mathematics, at grade 4, NAEP scores improved for all groups between 2003 and 2007, while the overall average score and that for U.S. non-Hispanic whites has changed little since then (Figure 10). However, the average scores of U.S. Asians, U.S. Hispanics, and U.S. non-Hispanic Blacks continued to rise through 2013, after which

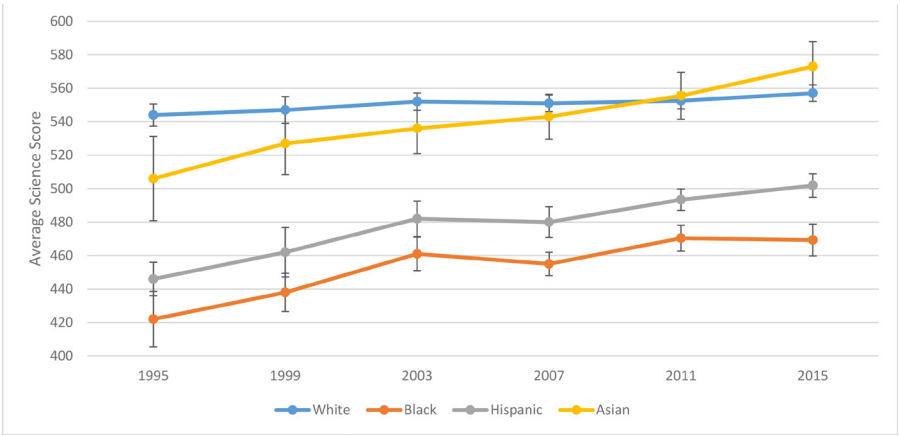

FIGURE 7. Average Science Scores in TIMSS: U.S. Eighth Graders by Race-Ethnicity, 1995–2015.

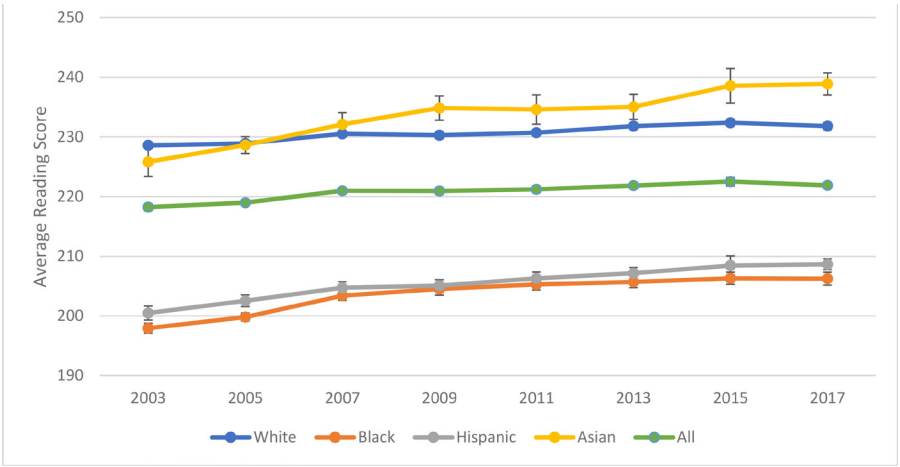

FIGURE 8. NAEP Reading Trends by Race-Ethnicity: Grade 4.

they may have declined slightly. The trends in math scores in the eighth grade are similar to those at the fourth-grade level, except the scores of U.S. Asians have continued to increase.

What is the overall story of these test score trends? In most cases, across the three large-scale assessments—PISA, TIMSS, and NAEP—minority scores have increased for the past several years, though by far less than would be required to eliminate the achievement gap between U.S. non-Hispanic whites and U.S. non-Hispanic Blacks or U.S. Hispanics. Given the favorable performance trends among the minority groups, why is there so little overall trend? There are two reasons. First, non-Hispanic whites remain the majority group among school-age

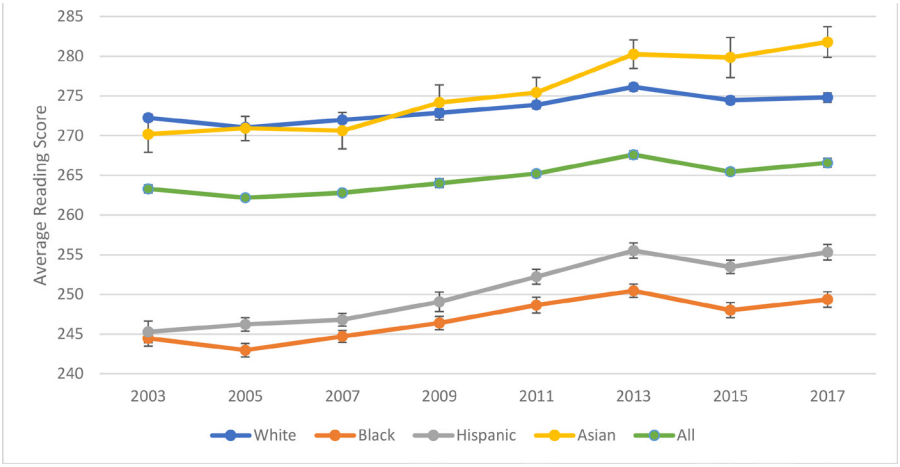

FIGURE 9. NAEP Reading Trends by Race-Ethnicity: Grade 8.

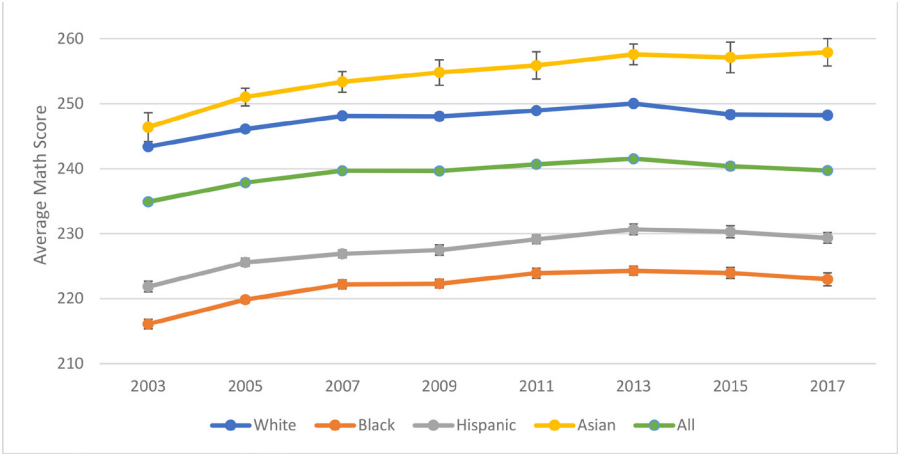

FIGURE 10. NAEP Math Trends by Race-Ethnicity: Grade 4.

youth in the United States and their performance has not varied greatly. But, second, the share of non-Hispanic whites among American students is declining, while that of minorities—especially Hispanics—is increasing. To illustrate this point, Figure 12 shows trends in the share of each racial-ethnic group among U.S. 15-year-olds—the target population of PISA—between 1997 and 2030. While the last 12 years of this series may appear to be a projection, it is not. All of these past, current, and future 15-year-olds have already been born. In sum, there is a push and pull between observed changes in the academic performance of minorities and in their growing share of the population. The former tends to raise overall average academic achievement, while the

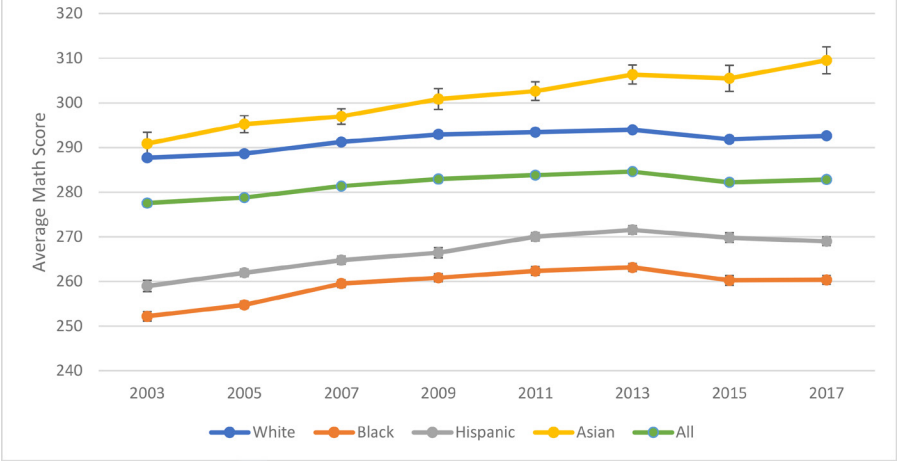

FIGURE 11. NAEP Math Trends by Race-Ethnicity: Grade 8.

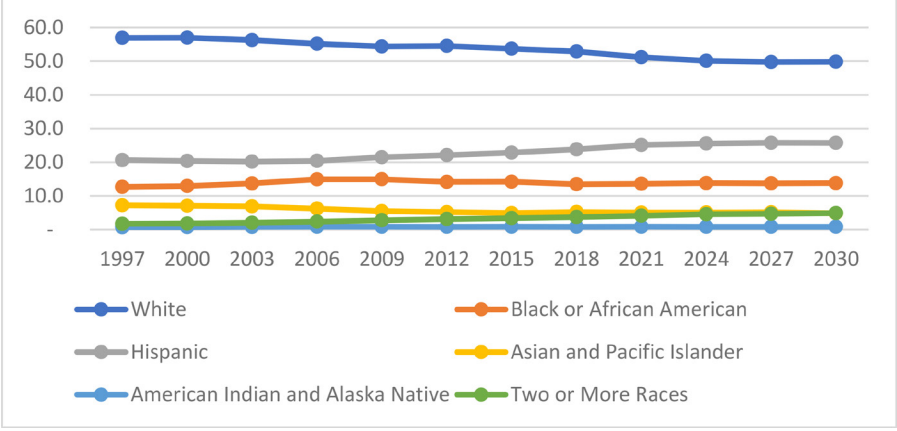

FIGURE 12. Trends in Percentage of 15-Year-Olds by Race-Ethnicity, 1997–2030.

latter—assuming continued performance differences among groups—will tend to lower it.

What Can Be Done

The tug-of-war between achievement differentials and population composition will continue in the absence of changes in the achievement gap, for there will be no pause in cohort replacement. Thus, for years to come the trajectory of overall achievement levels in the United States depends on possible changes in the achievement gap. Is there a proof of concept that the middling performance of the United States in international comparisons could be reduced substantially by raising academic

performance among minorities? What would be required of such a proof? There are many claims of success. These include neighborhood demonstration zones, no-excuses charter schools, and a variety of other school, teacher, and classroom policies and practices. While some successes are well-established, the claims are sometimes short-lived or fail to account for selective school entry and school-leaving (for example, Hoxby 2004; Betts and Hill 2010; Toma and Zimmer 2012; Welner 2013; Gamoran and Fernandez 2018; and studies cited therein).

A persuasive case should be based on a reasonably large and heterogeneous population of students, sustained across time, covering every age and grade level, and based on sound, comparable, and widely-used assessments. Is there such a case?

For example, consider the following scenario (Smith 2012):

> If someone asked you to describe expected achievement scores in a student population where a) many [families] have high personal debt with only a single parent at home; b) 40% of the school population is Latino or black; and c) students can expect to change schools between six and nine times as they move through primary and secondary school, below average results would probably come to mind. All of these stressors, it would be fair to assume, could contribute to difficulty with math, reading and other school skills, setting students up for an uphill struggle in the classroom.

But in this case, there isn't "an uphill struggle in the classroom" (for a compelling personal account, see Burnett 2019). The passage above describes the social circumstances of students in schools operated by DoDEA, which serves about 80,000 children of military personnel in the United States and an almost equal number internationally. Fortunately, the academic performance of all DoDEA students has regularly been assessed by the NAEP and reported by grade level and race-ethnicity.

Table 9 shows the mean and standard deviation of reading achievement in 2017 in national public schools and in DoDEA schools for the four largest racial-ethnic groups in the fourth and eighth grades. Except among U.S. Asians and Pacific Islanders in the eighth grade, mean test scores are higher in the DoDEA schools than in the nation's public schools. In that one exceptional case—where the mean scores are higher than in every other group—the difference in test scores is negligible. Moreover, as indicated by the standard deviations, the variability in test scores is substantially lower in the DoDEA schools than in the nation's public schools, for all students combined and within every racial-ethnic subgroup at both grade levels. The standard deviations are only 76 to 83 percent as large in the DoDEA schools than in the nation. That is, not only is reading achievement almost always higher

Reading, Grade 4		Mean	Std. Error	Std. Dev.	Std. Error
All	National public	221	(0.2)	38	(0.1)
	DoDEA	234	(0.6)	29	(0.6)
White	National public	231	(0.3)	35	(0.2)
	DoDEA	240	(1.0)	28	(0.8)
Black	National public	205	(0.5)	36	(0.3)
	DoDEA	226	(1.9)	28	(1.6)
Hispanic	National public	208	(0.5)	38	(0.3)
	DoDEA	229	(1.4)	31	(1.2)
Asian/Pacific Islander	National public	238	(1.0)	37	(0.7)
	DoDEA	233	(2.9)	29	(2.6)
Reading, Grade 8					
All	National public	265	(0.3)	36	(0.1)
	DoDEA	280	(0.8)	28	(0.7)
White	National public	274	(0.3)	33	(0.2)
	DoDEA	284	(1.2)	26	(1.0)
Black	National public	248	(0.5)	34	(0.3)
	DoDEA	268	(2.1)	26	(1.2)
Hispanic	National public	255	(0.5)	34	(0.3)
	DoDEA	277	(1.8)	28	(1.6)
Asian/Pacific Islander	National public	281	(1.0)	37	(0.9)
	DoDEA	283	(3.1)	30	(2.3)

TABLE 9. Reading Achievement by Grade Level, NAEP 2017: National Public and DoDEA Students.
Source: U.S. Department of Education, Institute of Education Sciences, NCES, NAEP 2017 Reading Assessment.

in the DoDEA schools than in the nation's public schools, but test performance is substantially more equal in DoDEA schools, overall and among racial-ethnic groups.

Table 10 shows comparable statistics on achievement in mathematics in 2017 NAEP for DoDEA schools and national public schools. Again, except among U.S. Asians and Pacific Islanders (in both the fourth and eighth grades), mean achievement test scores are larger among DoDEA students than in the national student population. And, again, there is less variability—and greater equality—in test

performance among DoDEA students in every population group than in the national population. In mathematics, the standard deviations are 72 to 88 percent as large among DoDEA students as among all students.

How large and important are the mean differences in achievement test scores? Table 11 shows a rearrangement of the mean reading achievement scores that highlights both the performance differences between DoDEA and all public-school students and those between U.S. non-Hispanic whites and U.S. non-Hispanic Blacks and U.S. Hispanics. The bottom two rows of the table show differences in mean test scores between DoDEA and public-school students at each grade level. Note that the DoDEA advantage is smaller for U.S. non-Hispanic whites and negligible for U.S. Asians and Pacific Islanders. However, for both minority groups, the DoDEA schools have an advantage of about 20 test score points relative to all public schools at each grade level. How important is this? Note, as shown in Table 9, that the overall standard deviation of reading test scores is 38 in the fourth grade and 36 in the eighth grade. Thus, the achievement gap in reading test performance is less than half as large in DoDEA schools as in the nation's public schools. By any standard, that is a huge reduction in racial-ethnic inequality. Further, as shown in the four cells in the lower-right corner of Table 11, the DoDEA advantage is 10 to 12 points larger among U.S. non-Hispanic Blacks and U.S. Hispanics than among U.S. non-Hispanic whites and U.S. Asians and Pacific Islanders. That is, the DoDEA schools are most successful in reducing inequality of reading achievement among the traditionally lower-scoring groups. Table 12 displays similar differentials in mathematics achievement in 2017 NAEP. Again, the DoDEA advantage is greater among the lower-scoring groups and is close to half a standard deviation. And again, the DoDEA advantage is substantially larger among the two lower-scoring groups.

Not only did achievement test scores in DoDEA schools compare favorably with those in other American schools in 2017, as shown in Figures 13 and 14, those differentials have been consistent from 1998 onward, both in reading and mathematics and at grades 4 and 8. What is known about the differences between DoDEA and public schools? The superior academic performance, reduced inequality in test scores, and reduced minority-majority differentials of DoDEA schools have long been recognized (Anderson, Bracken, and Bracken 2000; Wright et al. 2000; Smrekar et al. 2001; Bridglall and Gordon 2003; Department of Defense Education Activity 2010). Several studies have identified resources and practices that may account for this success. For example, Smrekar et al. (2001, i) report that the factors accounting for high academic achievement include:

Mathematics, Grade 4		Mean	Std. Error	Std. Dev.	Std. Error
All	National public	239	(0.2)	32	(0.1)
	DoDEA	249	(0.5)	25	(0.4)
White	National public	248	(0.2)	29	(0.1)
	DoDEA	252	(0.8)	25	(0.7)
Black	National public	223	(0.5)	29	(0.3)
	DoDEA	239	(1.4)	25	(1.2)
Hispanic	National public	229	(0.4)	30	(0.2)
	DoDEA	244	(1.3)	24	(1.0)
Asian/Pacific Islander	National public	258	(1.1)	33	(0.6)
	DoDEA	254	(2.1)	24	(1.3)
Mathematics, Grade 8					
All	National public	282	(0.3)	39	(0.2)
	DoDEA	293	(0.7)	32	(0.6)
White	National public	292	(0.3)	36	(0.2)
	DoDEA	298	(1.2)	31	(0.9)
Black	National public	260	(0.5)	34	(0.3)
	DoDEA	277	(2.1)	30	(1.6)
Hispanic	National public	268	(0.5)	35	(0.3)
	DoDEA	287	(1.9)	30	(1.3)
Asian/Pacific Islander	National public	310	(1.5)	42	(0.8)
	DoDEA	303	(2.9)	34	(2.8)

TABLE 10. Mathematics Achievement by Grade Level, NAEP 2017: National Public and DoDEA Students.
Source: U.S. Department of Education, Institute of Education Sciences, NCES, NAEP 2017 Mathematics Assessment.

- Centralized direction-setting with local decision-making
- Policy coherence and regular data flow regarding instructional goals, assessments, accountability, and professional training and development
- Sufficient financial resources linked to instructionally relevant strategic goals

	White	Black	Hispanic	Asian	W-B	W-H
U.S. Public, Grade 4	231	205	208	238	26	23
DoDEA, Grade 4	240	226	229	233	14	11
U.S. Public, Grade 8	274	248	255	281	26	19
DoDEA, Grade 8	284	268	277	283	16	7
DoDEA - Public, Grade 4	9	21	21	-5	12	12
DoDEA - Public, Grade 8	10	20	22	2	10	12

TABLE 11. Mean Reading Achievement in 2017 NAEP by Race-Ethnicity: U.S. Public Schools and DoDEA Students.

	White	Black	Hispanic	Asian	W-B	W-H
U.S. Public, Grade 4	248	223	229	258	25	19
DoDEA, Grade 4	252	239	244	254	13	8
U.S. Public, Grade 8	292	260	268	310	32	24
DoDEA, Grade 8	298	277	287	303	21	11
DoDEA - Public, Grade 4	4	16	15	-4	12	11
DoDEA - Public, Grade 8	6	17	19	-7	11	13

TABLE 12. Mean Mathematics Achievement in 2017 NAEP by Race-Ethnicity: U.S. Public Schools and DoDEA Students.

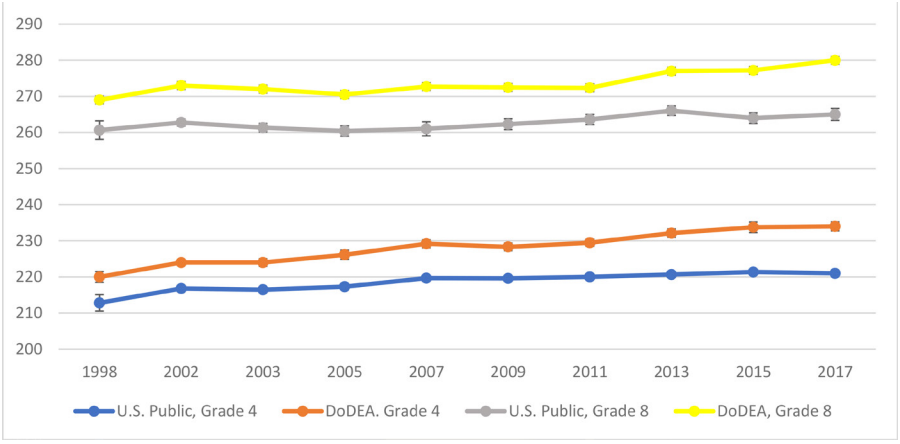

FIGURE 13. Trends in NAEP Reading Achievement: U.S. Public and DoDEA Schools.

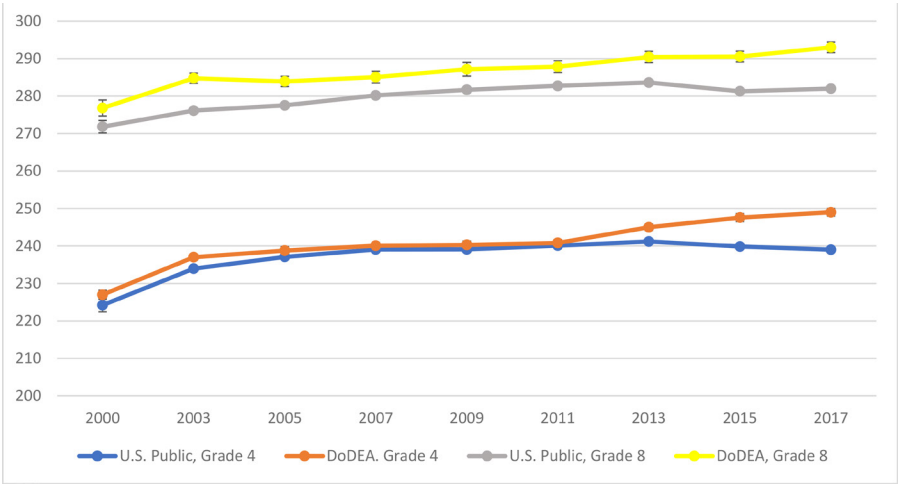

FIGURE 14. Trends in NAEP Math Achievement: U.S. Public and DoDEA Schools.

- Staff development that is job-embedded, intensive, sustained over time, relevant to school improvement goals, and linked to student performance
- Small school size, conducive to trust, communication, and sense of community
- Academic focus and high expectations for all students
- Continuity of care for children in high-quality preschools and after-school programs

- A "corporate commitment" to public education that is material and symbolic and that is visible and responsive to parents within the school community

Smith's (2012) journalistic account provides a richer description of the environment and practices of DoDEA schools:

- Housing security
- Routine health care services
- Support network of military families and personnel
- A culture of learning
- Hands-on accountability
- Not subject to NCLB
- Higher than average funding
- Highly organized, centralized chain of command
- Collaboration with unions
- Clear achievement goals
- Parental involvement and family resource centers
- Focus on good teaching
- Regular surveys of parents, teachers, and schools
- Wise, diagnostic use of tests

At the same time, that account mentions problems, including frequent moves, bullying, and inadequate services for disabled students.

The several accounts of DoDEA success leave an obvious question unanswered. Is it something special about teaching and learning in DoDEA schools? Or are student populations selected to achieve similarly and at high levels? Unfortunately, there is no compelling evidence about the effect of family background and community factors on the comparative success of students in DoDEA schools. For example, prompted by parental dissatisfaction with European DoDEA schools following troop drawdowns in Europe, the Institute for Defense

Analyses carried out an extensive study of academic performance and school organization throughout the DoDEA system (Anderson, Bracken, and Bracken 2000; Wright et al. 2000). The authors of the study were evidently aware of the effects of population composition on academic performance (Wright et al. 2000, II-11):

> In some cases, comparisons of DoDEA with other state or local systems should take account of the unusual demographics of DoDEA students, i.e.:
>
> • At least one, if not both, parents are employed.
>
> • All families have adequate housing, food, clothing, and medical care.
>
> • Families live in a relatively drug-free and low crime environment.
>
> • The military community is well educated and under stands the value of education.
>
> • The student population has a diverse cultural and ethnic background.
>
> • The annual mobility rate among students in the schools exceeds 35%.

However, there was no way of controlling these factors and others, e.g., parents' ability, education, or income, in the analysis. Most of the report consisted of comparisons of aggregate DoDEA data with that from other school systems, districts, and localities.

I have been able to locate only one intensive, comparative study of academic achievement in DoDEA schools and public schools, the Princeton University doctoral thesis of Leslie R. Hinkson (2007). It is based on data that include social and economic background characteristics of individual students and, also, characteristics of the schools they attended. Hinkson's analyses focus on Black-white differences in reading scores in the 1998 administration of the NAEP. The data come from a well-designed national sample of almost 63,000 Black or white students that covered 62 DoDEA schools. Unfortunately, Hinkson's analysis does not directly address the role of social background in Black-white achievement differences within DoDEA schools or between DoDEA schools and other public schools.

Hinkson cites Moskos (1989) in reporting that, relative to the general population, white military entrants are negatively selected for

educational attainment while Black entrants are positively selected. Indeed, Hinkson (2007, 35) writes, "Black enlisted personnel are slightly better educated and tend to enter into military service from higher income communities than their White counterparts." She adds, ". . . the difference in the distribution of AFQT scores by race within the military is much smaller than in the civilian population. This may help explain both why a racial test score gap exists between the children of Black and White military personnel but also why these gaps are smaller than those found in the general population" (35). Indeed, Moskos (1989, 78) wrote that "[T]oday the army's enlisted ranks are the only major segment of American society where the educational levels of non-Hispanic Blacks surpass those of whites." However, Moskos's data refer primarily to enlistees around 1980, aged 18–24, and it is problematic to extrapolate from them to the characteristics of military (and non-military) parents of students in DoDEA schools. Such a generalization would be hampered by the characteristics of commissioned officers, attrition from the military, selection into marriage and childbearing within the military, and specific military postings, as well as by the passage of time.

In fact, in the NAEP sample analyzed by Hinkson, and contrary to Moskos's data on military entrants, the parents of both white and Black students in the DoDEA schools had completed more education than the parents of students in public schools. Further, the educational attainment of the parents of white students in DoDEA schools exceeded that of the parents of Black students in DoDEA schools (Hinkson 2007, 91, Table 4c). That is, the 1998 NAEP data about parents provided no support for Hinkson's inferences about educational selectivity in DoDEA schools that were based on Moskos's essay.

Moreover, Hinkson's (2007, chapters 4 and 5) analyses do not directly address the degree to which selectivity accounts for the substantial convergence of Black and white test scores in DoDEA schools, nor do they explain why the performance of white students in DoDEA schools is comparable to that of white students in public schools. Rather, they consist of separate analyses of test scores in public, DoDEA, and Catholic schools and the extent to which background and school characteristics differentially affect academic performance within each sector. That is, Hinkson's analyses do not explain differences in reading achievement between military and civilian sectors.

The fact remains that there is no conclusive evidence about the role of a supportive social environment and specific educational practices in the success of DoDEA schools as compared to the role of social selection into the military. It should be a high priority to find out which among their distinct characteristics and practices account for their

extraordinary success. However, there is a complementary account of similarly effective educational practices, in the form of an evaluation of "community schools." Maier et al. (2017) examined some 143 recent, high-quality research studies of school effectiveness. Curiously, that report made no mention of the DoDEA schools or studies of them. The findings of the study were encapsulated in four "pillars" of effective education (16):

> • **Integrated student supports** address out-of-school barriers to learning through partnerships with social and health service agencies and providers, ideally coordinated by a dedicated professional staff member. Some employ social-emotional learning, conflict resolution training, trauma-informed care, and restorative justice practices to support mental health and lessen conflict, bullying, and punitive disciplinary actions, such as suspensions.
>
> • **Expanded learning time and opportunities**, including after-school, weekend, and summer programs, provide additional academic instruction, individualized academic support, enrichment activities, and learning opportunities that emphasize real-world learning and community problem solving.
>
> • **Family and community engagement** brings parents and other community members into the school as partners with shared decision-making power in children's education. Such engagement also makes the school a neighborhood hub providing adults with educational opportunities, such as ESL classes, green card or citizenship preparation, computer skills, art, STEM, etc.
>
> • **Collaborative leadership and practice** build a culture of professional learning, collective trust, and shared responsibility using such strategies as site-based leadership/governance teams, teacher learning communities, and a community school coordinator who manages the complex joint work of multiple school and community organizations.

There is marked similarity between these broad features of community and school structure and practice with the characteristics of DoDEA schools, as described above. Yet only 24 of the 143 sites in the community school study had established all four of the "pillars." That is, there are few exemplars, outside of the DoDEA schools, of the wraparound package of their environment and services. In sum, there is reason for hope, but no conclusive evidence that the multiple, positive features of DoDEA schools—or other community schools—could be introduced and succeed throughout American society.

To return to the opening theme of this essay, racial-ethnic achievement gaps are the only obstacles that stand between present levels of academic achievement among American students and world-class performance. While achievement gaps are narrowing—and those between the high-achieving groups and U.S. Hispanics can be explained to a substantial degree by differences in socioeconomic origins—overall improvement in achievement test scores is slowed by changes in the racial-ethnic composition of student populations. The good news, exemplified by DoDEA schools, is that present knowledge could yield a much faster narrowing of achievement gaps. The real question is whether or when America's political and educational leaders will marshal the necessary will and resources to eliminate the achievement gap.

References

Anderson, Lowell Bruce, Jerome Bracken, and Marilyn C Bracken. 2000. *Review of Department of Defense Education Activity (DoDEA) Schools. Volume II: Quantitative Analysis of Educational Quality*. Alexandria, VA: Institute for Defense Analyses.

Betts, Julian R., and Paul T. Hill, eds. 2010. *Taking Measure of Charter Schools: Better Assessments, Better Policymaking, Better Schools*. Plymouth, UK: Rowman & Littlefield Education.

Bracey, Gerald W. 1995. "The Assessor Assessed: A 'Revisionist' Looks at a Critique of the Sandia Report." *Journal of Educational Research* 88, no. 3: 136–44.

Bridglall, Beatrice L., and Edmund W. Gordon. 2003. "Raising Minority Academic Achievement: The Department of Defense Model. Pedagogical Inquiry and Praxis." *Pedagogical Inquiry and Praxis* 5.

Burnette, Daarel, II. 2019. "The Black Achievement Paradox Nobody's Talking About." *EducationWeek*. Last modified January 9, 2019. https://www.edweek.org/ew/articles/2019/01/09/the-black-achievement-paradox-nobodys-talking-about.html?intc=EW-BIG-NXT.

Carnoy, Martin, and Richard Rothstein. 2013. *What Do International Tests Really Show about U.S. Student Performance?* Washington, DC: Economic Policy Institute.

Carson, C. C., R. M. Huelskamp, and T. D. Woodall. 1993. "Perspectives on Education in America: An Annotated Briefing." *Journal of Educational Research* 86, no. 5: 259. doi: 10.1080/00220671.1993.9941211.

Department of Defense Education Activity (DoDEA). 2010. *Annual Report, 2010*. Washington, DC: U.S. Department of Defense.

Gamoran, Adam, and Cristina M. Fernandez. 2018. "Do Charter Schools Strengthen Education in High-Poverty Urban Districts?" In *Choosing Charters: Better Schools or More Segregation?*, edited by Iris C. Rotberg and Joshua L. Glazer, 133–52. New York: Teachers College Press.

Hanushek, Eric A., Paul E. Peterson, and Ludger Woessmann. 2012. *Achievement Growth: International and US State Trends in Student Performance*. PEPG Report No. 12-03. Program on Education Policy and Governance, Harvard University.

———. 2010. *US Math Performance in Global Perspective: How Well Does Each State Do at Producing High-Achieving Students?* PEPG Report No. 10-19. Program on Education Policy and Governance, Harvard University.

Hinkson, Leslie R. 2007. "Schools of the Nation: Department of Defense Schools and the Black-White Test Score Gap." Unpublished doctoral thesis, Department of Sociology, Princeton University.

Hoxby, Caroline Minter. 2004. *A Straightforward Comparison of Charter Schools and Regular Public Schools in the United States*. National Bureau of Economic Research.

Huelskamp, Robert M. 1993. "Perspectives on Education in America." *Phi Delta Kappan* 74, no. 9: 718–21.

Institute of Education Sciences (IES), and National Center for Education Statistics (NCES). 2016. "Welcome to PISA 2015 Results." Accessed February 13, 2017.

Maier, Anna, Julia Daniel, Jeannie Oakes, and Lilvia Lam. 2017. *Community Schools as an Effective School Improvement Strategy: A Review of the Evidence*. Palo Alto, CA: Learning Policy Institute.

Moskos, Charles C. 1989. "The All-Volunteer Force and the Marketplace." In *Who Defends America: Race, Sex, and Class in the Armed Forces*, edited by Edwin Dorn. Washington, DC: Joint Center for Political Studies Press.

Organisation for Economic Cooperation and Development (OECD). 2016. *PISA 2015 Results. Volume I: Excellence and Equity in Education*. Paris: OECD Publishing, Programme for International Student Assessment (PISA).

Peterson, Paul E., Ludger Woessmann, Eric A. Hanushek, and Carlos X. Lastra-Anadón. 2011. *Globally Challenged: Are US Students Ready to Compete? The Latest on Each State's International Standing in Math and Reading*. PEPG 11-03. Program on Education Policy and Governance, Harvard University.

Resmovits, Joy. 2016. "American Teens' Math Scores Fall on an International Test." *Los Angeles Times*, December 6, 2016.

Smith, S. E. 2012. "Why Are Students at Military Base Schools Out-achieving Their Civilian Peers?" *AlterNet*. https://www.alternet.org/education/why-are-students-military-base-schools-out-achieving-their-civilian-peers.

Smrekar, Claire, James W. Guthrie, Debra E. Owens, and Pearl G. Sims. 2001. *March toward Excellence: School Success and Minority Student Achievement in Department of Defense Schools. A Report to the National Education Goals Panel. Lessons from the States*. Washington, DC: National Educational Goals Panel.

Stedman, Lawrence C. 1997. "International Achievement Differences: An Assessment of a New Perspective." *Educational Researcher* 26, no. 3: 4–15. doi: 10.3102/0013189X026003004.

———. "The Sandia Report and U.S. Achievement: An Assessment." *Journal of Educational Research* 87, no. 3: 133–46. doi: 10.1080/00220671.1994.9941235.

Stedman, Lawrence C., Ina V. S. Mullis, and Michael Timpane. 1998. "An Assessment of the Contemporary Debate over U.S. Achievement." *Brookings Papers on Education Policy* 1: 53–121.

Strauss, Valerie. 2017. "DeVos: Outcomes at U.S. Schools Are So Bad, They Probably Can't Get Worse." *Washington Post*, March 29, 2017.

Toma, Eugenia, and Ron Zimmer. 2012. "Two Decades of Charter Schools: Expectations, Reality, and the Future." *Economics of Education Review* 31, no. 2: 209–12.

Tucker, Marc. 2016. "The 2015 PISA Results: What Do They Mean." *National Center on Education and the Economy*. http://ncee.org/2016/12/the-2015-pisa-results-what-do-they-mean/.

U.S. Department of Education, National Commission on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform*. Washington, DC: National Commission on Excellence in Education.

Welner, Kevin G. 2013. "The Dirty Dozen: How Charter Schools Influence Student Enrollment." *Teachers College Record*.