

RESEARCH REPORT
OCCASIONAL PAPER

US Educational Achievement on International Assessments: The Role of Race and Ethnicity

Ben Dalton

May 2011

RTI Press

RTI
INTERNATIONAL

About the Author

Ben Dalton, PhD, is a research education analyst in RTI International's Education Studies Division.

RTI Press publication OP-0005-1105

This PDF document was made available from www.rti.org as a public service of RTI International. More information about RTI Press can be found at <http://www.rti.org/rtipress>.

RTI International is an independent, nonprofit research organization dedicated to improving the human condition by turning knowledge into practice. The RTI Press mission is to disseminate information about RTI research, analytic tools, and technical expertise to a national and international audience. RTI Press publications are peer-reviewed by at least two independent substantive experts and one or more Press editors.

Suggested Citation

Dalton, B. US Educational Achievement on International Assessments: The Role of Race and Ethnicity. RTI Press publication No. OP-0005-1105. Research Triangle Park, NC: RTI Press. Retrieved [date] from <http://www.rti.org/rtipress>.

This publication is part of the RTI Research Report series. Occasional Papers are scholarly essays on policy, methods, or other topics relevant to RTI areas of research or technical focus.

RTI International
3040 Cornwallis Road
PO Box 12194
Research Triangle Park, NC
27709-2194 USA

Tel: +1.919.541.6000
Fax: +1.919.541.5985
E-mail: rtipress@rti.org
Web site: www.rti.org

©2011 Research Triangle Institute. RTI International is a trade name of Research Triangle Institute.

All rights reserved. Please note that this document is copyrighted and credit must be provided to the authors and source of the document when you quote from it. You must not sell the document or make a profit from reproducing it.

[doi:10.3768/rtipress.2011.op.0005.1105](https://doi.org/10.3768/rtipress.2011.op.0005.1105)

www.rti.org/rtipress

US Educational Achievement on International Assessments: The Role of Race and Ethnicity

Ben Dalton

Abstract

The debate about the performance of US students on international assessments of educational achievement routinely fails to account for one consistently stark result: US achievement is bifurcated between a group of high-performing Asian and white students and an exceptionally low-performing group of black and Hispanic students. By summarizing results across 20 major international tests conducted since 1995, this research paper shows that when US racial and ethnic groups are separately compared with other countries, Asian and white students regularly perform at or near the top of international rankings, while black and Hispanic students typically rank at or near the bottom. Furthermore, the United States has a substantially larger minority population than all other developed countries, and minority status is not synonymous with internationally comparable factors such as socioeconomic level or immigrant status. The fact that overall US scores are disproportionately influenced by race and ethnicity suggests that researchers and reformers must reconsider how they use and interpret the results of international achievement tests. The research paper recommends providing and analyzing cross-national data separately by ethnic groups to inform approaches to improving education for all students.

Contents

Introduction	2
Background: Assessment Programs and Comparison Countries	2
US Racial and Ethnic Group Performance in an International Context	4
Minority Achievement Gaps in Comparison Countries	7
Accounting for Racial and Ethnic Differences	10
Implications and Research Recommendations	12
Conclusion	15
References	15
Acknowledgments	Inside back cover

Introduction

The performance of US students on international assessments of educational achievement has been subject to a variety of interpretations. Media and policy organizations commonly pronounce US performance lackluster (Council on Competitiveness, 2005; Hanushek, Peterson, & Woessmann, 2010; National Academy of Science, 2007), while many researchers argue that US performance is not as poor as commonly believed or that it is even above average (Boe & Shin, 2005; Koretz, 2009; Lowell & Salzman, 2007). However, the debate about the overall performance of US students misses one of the starkest and most consistent findings to emerge from international assessments: US educational achievement is bifurcated between a group of high-performing Asian and white students and an exceptionally low-performing group of black and Hispanic students. If US racial and ethnic groups were separately compared with other countries, Asian and white students would regularly perform at or near the top of international rankings, while black and Hispanic students would routinely rank at or near the bottom.

Though racial and ethnic differences in achievement are well known in the United States (Jencks & Phillips, 1998; Kao & Thompson, 2003) and regularly noted in reports describing US achievement in an international context (e.g., Gonzales et al., 2008), the relatively large contribution racial and ethnic differences make to overall US performance has not been widely recognized or discussed. Furthermore, the implications of these differences for educational research and policy have been little explored. In this paper, I aim to focus attention on the unique role that race and ethnicity play in US performance on international assessments by (a) systematically examining how US racial and ethnic groups perform compared with other developed countries, (b) comparing racial and ethnic differences in the United States to evidence of similar differences elsewhere, (c) comparing the size of the US minority population to the size of minority populations in other nations, and (d) discussing the lack of comprehensive explanations for racial and ethnic differences in achievement. The paper concludes by providing two research recommendations for

improving how international assessment programs inform educational policy and practice in the United States.

Background: Assessment Programs and Comparison Countries

Though cross-national educational research has origins as far back as the 1960s, it was the Third International Mathematics and Science Study in 1995 that launched a boom in comparative research, particularly in math and science, and established the generally consistent methodology and reporting practices that subsequent assessments have followed. Later rechristened as a series of studies (the Trends in International Mathematics and Science Study [TIMSS]), TIMSS has surveyed an increasing number of countries in studies of math and science conducted in 1995, 1999, 2003, and 2007 (with another planned for 2011). Another program, the Progress in International Reading Literacy Study (PIRLS), has conducted two reading assessments (2001 and 2006) and is planning for more in the future (2011). TIMSS and PIRLS are both conducted under the aegis of the International Association for the Evaluation of Educational Achievement (IEA).

TIMSS and PIRLS are complemented by the Program for International Student Assessment (PISA), which has been conducted in 2000, 2003, 2006, and 2009.¹ Unlike TIMSS and PIRLS, PISA rotates its main subject focus among math, science, and reading, although it does some additional testing in the subjects that are not a main focus in a given year. PISA is conducted under the authority of the Organisation for Economic Co-operation and Development (OECD), a collection of 20 democratic, industrialized countries. The National Center for Education Statistics (NCES) of the US Department of Education conducts the US data collections and produces official US reports for all three programs.

¹ An international survey of adult skills and competencies, the Program for the International Assessment of Adult Competencies [PIAAC], is also planned for 2011. Another survey, the Adult Literacy and Lifeskills Survey of 2003, was conducted with a nonschool population and is not reviewed here. Finally, the Civic Education Study of 1999, though focused on a school-age population, has been conducted only once and is not included here because its results rely on a single, unique set of participating countries and national samples.

In TIMSS, assessments are designed to cover math and science topics that are typically included in the formal school curriculum. In contrast, in PIRLS and PISA, assessments are designed to gauge the overall “literacy” or familiarity with more general concepts and skills related to each subject—the kind of knowledge that can be gained outside of school as often as in it. Consistent with the curricular focus, TIMSS surveys students within targeted grades—specifically, students in the fourth and eighth years of compulsory schooling (corresponding to 4th grade and 8th grade in the United States)²—while PIRLS and PISA, with broader definitions of learning, sample whole ages, regardless of grade: age 9 in PIRLS and age 15 in PISA.

Both developed and developing countries have participated in these studies, and their participation has varied over time (the United States participated in all of the studies). From the perspective of the United States, however, only certain comparisons are reasonable. It makes little sense to compare the United States with economically less developed countries, many of whom have far lower school enrollment and attendance rates among their school-aged populations (this is even true for OECD members Mexico and Turkey), in addition to substantially lower incomes and less advanced infrastructure.

In addition, some have argued that very small developed countries or jurisdictions are unreasonable to compare with a large and diverse country such as the United States (Boe & Shin, 2005). The argument is that city-states, such as Singapore and the Hong Kong Special Administrative Region, and very small countries, such as Iceland and Luxembourg, do not have the kind of diversified economy or the overall population size to be appropriate comparisons (the latter two countries have less than half a million people—approximately equivalent to a mid-sized US metropolitan area). Another set of countries whose comparability to the United States has been

questioned includes more recent entrants to the community of developed capitalist economies, such as the Slovak Republic, the Czech Republic, Poland, and South Korea.

Though the debate about appropriate comparison countries is not settled (indeed, appropriate comparison groups may legitimately vary depending on research goals), I follow the practice of PISA and prior research (Boe & Shin, 2005) in limiting comparisons to member countries of the OECD. This restriction automatically excludes developing countries as well as unique city-states such as Singapore and Hong Kong, though it includes small countries like Iceland and Luxembourg.

However, since the composition of the OECD has shifted over time, I only include countries that have been part of the OECD for the full time period covered by these assessments (1995 to 2010). This criterion excludes more recent OECD members that may be legitimate comparison countries for a part of the time period—for example, the Slovak Republic, which joined the OECD in 2000. The most objectionable exclusions may be countries that joined the OECD near the beginning of the time period covered here: the Czech Republic (joined the OECD in December 1995), Hungary (May 1996), Poland (November 1996), and South Korea (December 1996). However, the effect of these exclusions is negligible: the Czech Republic, Hungary, and Poland regularly score near the US average across assessments, while South Korea is the only country regularly scoring above the US average. Including these countries does not change the results presented here.

As a final measure, again following PISA practice (OECD, 2007), I exclude any country in which less than 80 percent of the secondary school population is enrolled in school—this adds OECD members Mexico (72 percent) and Turkey (66 percent) to the list of excluded countries (UNESCO, 2010).

The final tally is 22 developed countries whose size, history, and level of economic development make them the most significant economic competitors to and partners with the United States: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan,

² In 1995, TIMSS also surveyed students in the final year of secondary school; however, there were significant concerns about the cross-national comparability of these samples, given the differences in the length of secondary education across national education systems, and this sample of students has not been assessed since. These results are not part of the current analysis.

Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom.³ International averages based on these countries provide a more consistent and appropriate basis of comparison with the United States than the “international” averages reported as part of TIMSS and PIRLS, which include developing countries, and the OECD average reported as part of PISA, which includes smaller countries and jurisdictions as well as some less developed countries.⁴ Nevertheless, even within this group, it may be worth noting, as Smith (2002) does, that some US states “are closer in character to many TIMSS [or PIRLS or PISA] nations than the United States itself” (p. 301).

US Racial and Ethnic Group Performance in an International Context

Racial and ethnic differences in achievement—specifically, lower achievement levels of blacks and Hispanics compared with their Asian and white peers—have been thoroughly documented in the American educational literature (Jencks & Phillips, 1998; Kao & Thompson, 2003; Sakamoto, Goyette, & Kim, 2010). Researchers have struggled to account for persistent differences between blacks and whites across multiple subjects and at all levels of education, and the last reauthorization of the federal Elementary and Secondary Education Act (i.e., No Child Left Behind) famously intended to reduce this and similar gaps. The relevancy of US gaps has not escaped the attention of those looking at US results from a cross-national perspective either. Boe and Shin (2005), Lowell and Salzman (2007), and prior researchers (Bracey, 1996; Westbury, 1993) have all noted the challenges for US education that these differences have created. Likewise, official reports

using international assessment data from NCES regularly indicate differences within the United States by race and ethnicity (Baer, Baldi, Ayotte, & Green, 2007; Baldi, Ying Jin, Skinner, Green, & Herget, 2007; Gonzales, 2004, 2008; Lemke et al., 2004; Ogle et al., 2003). However, no article or report has focused on race and ethnicity as a key marker of US achievement, and no research has undertaken to collect and summarize US results by race and ethnicity across multiple administrations of the international assessments.

In the absence of this information, researchers and commentators continue to make broad, overarching statements about US performance on international assessments (Darling-Hammond, 2010; Editorial: Our View on Education, 2010). The average US score is deemed representative of the entire US population; the diversity recognized in the reporting of domestic assessment results (such as from the National Assessment of Education Progress) is ignored in the face of simple international rankings of country averages. This tendency is misleading and results in policy recommendations that are at variance with the unique pattern of achievement among US students. For example, the influential Bill & Melinda Gates Foundation (2011) lists the following as one of their “top 10 facts about college readiness” in the United States: “On the 2006 PISA test, 15-year-old students in 23 countries scored higher than those in the United States in math literacy and 16 countries scored higher in science literacy.”⁵ Although the Bill & Melinda Gates Foundation recognizes the need to focus on low-income and minority students, its policy proposals appear to take such overarching statements at face value by pushing for comprehensive school reform models and overhauls of teacher evaluation systems.

In contrast, comparing the scores of US racial and ethnic groups with the average scores of other countries reveals a very different picture of US performance than that assumed by either side in the debate over the United States’ international standing.

³ Results for the United Kingdom are reported separately for England and Scotland in TIMSS and PIRLS. At times, results for subregions of countries such as Canada and Belgium have been reported separately in TIMSS and PIRLS. For the purposes of computing developed-country averages in the current analysis, these regions are included as their own jurisdictions.

⁴ An alternative grouping uses countries designated as “advanced economies” by the International Monetary Fund (e.g., see IMF, 2010, Table 4.5). This list includes very small countries such as Luxembourg and Malta. It also includes countries such as South Korea, the Czech Republic, and Slovenia.

⁵ This statement is misleading because it includes developing countries where substantial proportions of the 15-year-old population are not enrolled in school (and therefore not tested).

To show the different picture of US performance, I first present differences between the international average scale scores and the average scores of US students in each racial and ethnic group. Table 1 presents differences in scale scores for 20 separate tests conducted by the three major international assessment programs since 1995.⁶ Because average scores for certain minority groups (including what NCES refers to as Indian Americans, Alaska Natives, Native Hawaiians, and Pacific Islanders, as well as students reporting more than one race) are often either not reported or are grouped under the category “other,” results for only Asian, black, Hispanic, and white students are shown. Students of any race who report Hispanic ethnicity were included in the Hispanic category (by implication, all other groups are non-Hispanic). Students with insufficient English language experience were excluded from the study samples (this is the guideline for all countries that participate).

The most striking and consistent finding is that, in every major test across the 14 years covered by these assessments, blacks and Hispanics scored below the average developed country. Often, their scores were significantly lower: in the 1995 TIMSS assessments of math and science, black 8th graders scored over 102 points below the developed-country average. In the 2003 and 2006 PISA assessments (covering math and science, respectively), black 15-year-olds scored over 91 points below the developed-country average; blacks scored 60 points lower in the 2009 PISA assessment of reading. These results are about one full standard deviation below the scale score average, placing black students lower than *every single* developed country. Even in the six assessments in which black students did best relative to the developed-country average (TIMSS 2003 and 2007 4th-grade math and science tests, and PIRLS 2001 and 2006 reading literacy tests), black students ranked better than only students in Norway (in all cases) and the French portion of Belgium (in one case, PIRLS

2006). Black student performance on international assessments—on average, 61 points below the average developed country—is extremely poor.

Hispanic students’ performance, while generally not as poor as black students’ performance, was also consistently below the average developed country (by 40 points, on average) and consistently below other developed countries. Hispanic students recorded

Table 1. Differences in average scale scores between US racial and ethnic groups and the average developed country

Study, year, grade or age, and subject	Asian	Black	Hispanic	White
TIMSS 1995				
4th-grade math	—	−48	−12	36
4th-grade science	—	−60	−19	50
8th-grade math	—	−104	−80	−7
8th-grade science	—	−102	−78	20
TIMSS 1999 (4th grade not tested)				
8th-grade math	—	−85	−72	−4
8th-grade science	—	−89	−65	20
TIMSS 2003				
4th-grade math	37	−42	−22	28
4th-grade science	27	−30	−19	48
8th-grade math	—	−61	−44	16
8th-grade science	—	−56	−37	33
TIMSS 2007				
4th-grade math	67	−33	−11	35
4th-grade science	52	−33	−19	46
8th-grade math	48	−44	−26	32
8th-grade science	29	−59	−34	37
PIRLS 2001 (reading)	17	−32	−17	31
PIRLS 2006 (reading)	31	−33	−18	24
PISA 2000 (reading)	—	−63	−59	30
PISA 2003 (math)	−2	−91	−65	4
PISA 2006 (science)	−9	−99	−69	15
PISA 2009 (reading)	40	−60	−35	24
Average across all assessments	31	−61	−40	26

— Not available.

Note: The average developed-country scale score is based on Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. Results for US American Indians/Alaska Natives, multiracial students, and other groups are not presented because of limited availability. PIRLS tests students aged 9 and PISA tests students aged 15.

Sources: Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS), under the aegis of the International Association for the Evaluation of Educational Achievement (IEA), and Program for International Student Assessment (PISA), under the authority of the Organisation for Economic Co-operation and Development.

⁶ Only results from the major assessment scales from each program are presented. That is, I include only the combined literacy scale from PIRLS (not individual subscales) and the main focus of each PISA administration. Though PISA assesses three subjects in each round, only one subject is a primary focus (i.e., involves substantially more test questions and tailored survey questions) and produces sufficiently accurate data to enable comparison of racial/ethnic group performance in the United States with the performance of other countries.

their worst performance among 8th graders in 1995 and in 1999 (in both TIMSS subjects in each year) and among 15-year-olds in 2003 and 2006 (in PISA tests of math and science, respectively)—largely the same assessments in which blacks performed poorly. In these assessments, Hispanics, like blacks, averaged a lower score than every other developed country. Hispanic students performed comparatively best on eight assessments—TIMSS 4th-grade math and science tests in 1995, 2003, and 2007 and PIRLS reading tests in 2001 and 2006. In contrast to black students' performance, the Hispanic score on any given assessment was typically higher than two or three other developed countries, though still among the bottom ranks (and last in 4th-grade science in 1995).

In contrast to black and Hispanic student achievement, US Asian and white student achievement was typically above average or not statistically different from the average among developed countries. US Asian students, averaging 31 points above the developed-country averages, scored higher than the developed-country average on 9 of the 11 tests for which their scores were reported. Whites scored above average on 17 of the 20 tests (the difference for PISA 2003 was not statistically significant). Further, Asian and white students did not just score somewhat above average—in a mirror image of the poor performance of black and Hispanic students, US Asian and white students on occasion scored higher than *every single* developed country. This is true, for example, of US Asian students' 4th-grade math and science scores in 2007 (TIMSS) (and only below Japan in the 8th grade of that year's assessments) and of whites' 4th-grade science scores in 1995, 2003, and 2007 (TIMSS) (and sharing the highest-scoring spot with Japan on 2003's 8th-grade science test).

The truly divided nature of educational achievement in the United States can be seen in other ways as well. For example, Table 2 summarizes the ranking of each US racial and ethnic group within the list of developed countries or jurisdictions.⁷ In 14 of the 20 tests (70 percent of the time), blacks placed below the average

of every developed country; in 5 of the 6 remaining tests, blacks were second to last. Hispanics placed last in 10 of the 20 tests (50 percent of tests); in 4 of the remaining 10 tests, Hispanics are second to last; and they ranked poorly on all others except the TIMSS

Table 2. Ranking of US racial and ethnic group test performance among developed countries

Study, year, grade or age, and subject	Asian	Black	Hispanic	White	Number of countries
TIMSS 1995					
4th-grade math	—	8	5	3	8
4th-grade science	—	8	8	1	8
8th-grade math	—	9	9	5	9
8th-grade science	—	9	9	3	9
TIMSS 1999 (4th grade not tested)					
8th-grade math	—	6	6	4	6
8th-grade science	—	6	6	2	6
TIMSS 2003					
4th-grade math	2	9	8	3	10
4th-grade science	1	9	9	1	10
8th-grade math	—	10	9	4	10
8th-grade science	—	10	10	1	10
TIMSS 2007					
4th-grade math	1	12	9	2	13
4th-grade science	1	12	11	1	13
8th-grade math	2	8	7	2	8
8th-grade science	2	8	8	2	8
PIRLS 2001 (reading)	4	12	11	1	13
PIRLS 2006 (reading)	1	20	18	1	22
PISA 2000 (reading)	—	21	21	2	22
PISA 2003 (math)	13	22	22	11	22
PISA 2006 (science)	14	22	22	7	22
PISA 2009 (reading)	1	22	22	2	22
Average across all assessments	4	12	11	3	12

— Not available.

Note: The average developed-country scale score is based on Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. Results for US American Indians/Alaska Natives, multiracial students, and other groups are not presented because of limited availability. Each racial/ethnic group was ranked separately among the participating OECD countries (i.e., all four groups are not ranked simultaneously, so Asian and white students, for example, can have the same ranking in a given assessment). PIRLS tests students aged 9 and PISA tests students aged 15.

Sources: Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS), under the aegis of the International Association for the Evaluation of Educational Achievement (IEA), and Program for International Student Assessment (PISA), under the authority of the Organisation for Economic Co-operation and Development.

⁷ Though it is often pointed out that rankings are not very helpful in interpreting the standing of any country on a given assessment (e.g., Huang, 2009; Raudenbush & Kim, 2002), they provide useful summaries across multiple tests.

1995 4th-grade math test (where they were average). In contrast, in 8 of the 11 tests (73 percent) for which averages were published, Asian students ranked first or second and ranked fourth on one other test; in two tests, Asians were average performers. In 12 of all 20 tests (60 percent), white students ranked first or second, and ranked third another three times; in five tests, whites were average performers.

In other words, US black and Hispanic students consistently ranked at the bottom of cross-national achievement listings, while Asian and white students often ranked at or near the top. These rankings may come as a surprise to some, given the generic concern stimulated by reports of US achievement on international assessments (e.g., see Editorial: Our View on Education, 2010). The United States is simply not characterized by poor, average, or even above-average achievement; rather, it is characterized by above-average performance among Asians and whites and exceedingly low performance among blacks and Hispanics. Together, this bifurcated situation produces the middling picture that emerges from examining overall US scores and implies that something other than targeting the average student or indicting the entire educational system is needed. Before addressing the implications of these results, however, I turn to some possible objections and alternative explanations of the results' meaning.

Minority Achievement Gaps in Comparison Countries

One objection to the findings presented above is that comparing disadvantaged minority students in the United States with the average student in other developed countries is inherently unfair; a more apt comparison is to other minority students in those countries. Using a few countries that identify their minority students, we can compare gaps in the United States to gaps elsewhere, and this comparison provides clues about whether racial and ethnic gaps in the United States are out of the ordinary. However, such comparisons do not change the fact that the United States has a substantially larger proportion of minority students than nearly every other developed country. Even if the United States had narrower achievement gaps, the sheer size of its minority

population means that these groups contribute to the US average to a much greater extent than in other countries.

The United States possesses a racial and ethnic history that is unlike most other developed countries it can reasonably be compared with: no other Western European country, nor Japan, for example, had an extensive system of slavery, segregation, and racial disenfranchisement, nor have most of our developed-country peers experienced the extent of immigration that is embedded in the US experience.⁸

Only a few developed countries have anything approaching the US experience: in particular, Canada, Australia, and New Zealand possess the most similar histories and current demographics. Each has a native population that was overwhelmed by large numbers of colonists and immigrants, each has experienced some influx of East and South Asian immigrants, and each is majority European in origin (these are also the only three developed countries to report more immigrant students than the United States). Of course, as with every country, each has unique situations: Canadian society includes French- and English-speaking majorities, depending on region, as well as indigenous minorities and some Asian and African minorities; Australia and New Zealand have had more substantial recent East Asian immigration; and New Zealand has a relatively large proportion of Pacific Islanders. Nevertheless, comparing the achievement gaps in these countries with those in the United States could provide insights into the relative importance of race and ethnicity for US performance.

Although Canada reports on English- and French-language-minority students living in the other majority's regions, it does not gather data on racial or ethnic makeup of its population, which would

⁸ Though many developed European countries have experienced heightened immigration in recent decades, most remain relatively more homogeneous than the United States. For example, the OECD noted that the percentage of 15-year-olds in 2006 who were first- or second-generation immigrants averaged 10 percent across the group of developed countries (excluding the United States) analyzed here; the US percentage was 15 percent (OECD, 2007, p. 114). Countries such as Finland, Iceland, and Japan have very few immigrant students at all, and every Western European country has a lower immigration population than the United States.

be a more appropriate basis of comparison with the United States (see, for example, Bussière, Knighton, & Pennock, 2007). Therefore, Canada is not examined here. However, Australia reports indigenous (native or aboriginal) status, and New Zealand reports on ethnicity across multiple categories. Using information from Australia and New Zealand, we can obtain a glimpse into how the US racial and ethnic gaps might compare if minority status were more widely identified in international assessments. Results are reported in Table 3.

Table 3. TIMSS 2007 achievement differences between majority and minority students in Australia, New Zealand, and the United States

Country and race/ethnicity	Grade 4		Grade 8	
	Math	Science	Math	Science
Australia				
Indigenous	431	441	431	447
Nonindigenous	522	533	501	519
Nonindigenous–indigenous gap	91	92	70	72
New Zealand				
Asian	546	529	—	—
European	510	528	—	—
Māori	453	459	—	—
Pasifika	427	431	—	—
Other	491	502	—	—
Asian–Māori gap	93	70	—	—
Asian–Pasifika gap	119	98	—	—
European–Māori gap	57	69	—	—
European–Pasifika gap	83	97	—	—
United States				
Asian	582	573	549	543
Black	482	488	457	455
Hispanic	504	502	475	480
White	550	567	533	551
Asian–black gap	100	85	92	88
Asian–Hispanic gap	78	71	74	63
White–black gap	68	79	76	96
White–Hispanic gap	46	65	58	71

— Not available.

Note: Results for US American Indians/Alaska Natives, multiracial students, and other groups are not presented because of limited availability.

Source: Trends in International Mathematics and Science Study (TIMSS), under the aegis of the International Association for the Evaluation of Educational Achievement (IEA).

Australia estimated its indigenous population as 5 percent of their year (i.e., grade) 4 and year 8 population in TIMSS 2007 (Thomsen, McKelvie, & Murnane, 2006; Thomsen, Wernert, Underwood, & Nicholas, 2008). The achievement gap reported then between indigenous and nonindigenous students in math was 91 points in year 4 and 70 points in year 8. In science, it was 92 points in year 4 and 72 points in year 8. Though NCES did not report TIMSS 2007 results for US American Indians, who may be the most comparable group to Australian indigenous students, the indigenous–nonindigenous gap in science for Australia is similar to the US gap in reading between white and American Indian 9-year-olds in PIRLS 2006 (92 points). However, the 72-point Australian gap among somewhat younger year 8 students was smaller than the white–American Indian gap in science among 15-year-olds reported in PISA 2006 (87 points). Otherwise, the Australian gaps fall within the range of gaps for the United States across a variety of differences.

Only New Zealand collects and reports data in more detail. New Zealand collects data by five ethnic classifications and reported as part of TIMSS 2007 that 39 percent of “Year 5” students are in one of the minority groups:

Five broad ethnic classifications are used to describe ethnicity in New Zealand. They are: Pākehā/European, Māori, Pasifika, Asian, and “Other” ethnic groupings. The majority of Year 5 students in New Zealand were identified by their schools as Pākehā/European (61%) or Māori (19%). Pasifika (10%) and Asian (7%) students made up most of the rest of the ethnic groupings, with four percent of students categorised in the Other ethnic grouping. (Caygill & Kirkham, 2008, p. 29)

As Table 3 shows, the gaps among high- and low-achieving New Zealand ethnic groups are similar to the gaps observed among American racial and ethnic groups, with gaps running, for the most part, between 60 and 100 scale score points (Caygill, 2008; Caygill & Kirkham, 2008). Likewise, the pattern of findings in New Zealand mirrors the similar ethnic contrasts in the United States: Asian and European (white) New Zealand year 5 students scored higher than Māori and Pasifika students (8th grade–equivalent students were not tested in New Zealand in 2007). Though any specific comparison of gaps between the United

States and Australia or New Zealand is imperfect, the evidence here and from Australia suggests that the United States may not be particularly distinct in terms of the levels of achievement of its disadvantaged minorities. However, without data from additional countries that distinguish ethnicity or minority status from language use and immigrant status, these conclusions must be interpreted with caution (see the **Implications** section for more discussion on this point).

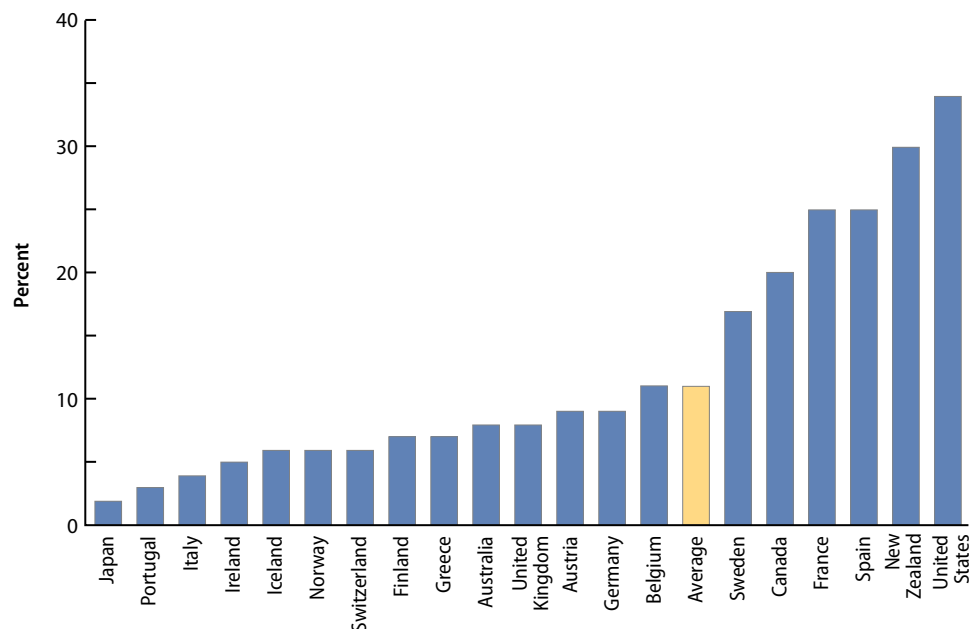
What can be asserted with more confidence is that the proportion of the US population that is minority is larger than the proportion of all other OECD countries' minority populations and often much larger. Estimates of cross-national minority populations are difficult to find, but one source has served as a source for data for a number of research projects: the Central Intelligence Agency (CIA) *World Factbook*, which reports the percentage of the adult population in ethnic groups for many countries (CIA, 2010). These are estimates of the adult population and likely underestimate ethnic proportions among school-aged students for most countries. For example, 41 percent of US 15-year-olds in PISA 2006 were nonwhite, while the CIA reports that 34 percent

of the US population is nonwhite. In the United States, at least, this difference results from higher rates of childbirth and the immigration of young families among minority groups. Nevertheless, these data provide a rough guide to the representation of ethnic groups within countries. Figure 1 presents the estimates for the developed countries examined in this analysis.

The United States has a greater proportion of minorities than *every other* developed country: 34 percent of the United States is nonwhite, compared with 11 percent nonmajority in the average developed country (and 8 percent average, taking out the United States). Some countries (such as France, Spain, and New Zealand) have large ethnic minorities, but no other country has as large a proportion of their population composed of minority groups as the United States. Half of these developed countries have less than 8 percent minority population—less than one-quarter of the United States' minority percentage.

Therefore, despite the suggestive evidence that US minority students perform no differently than ethnic minority students elsewhere, the relatively higher minority population in the United States influences the overall US performance to a much greater extent

Figure 1. Estimates of percentage of adult population that is ethnic minority



Note: Estimate for Denmark not available.

Source: Central Intelligence Agency. (2010). *The World Factbook 2010*.

than is possible in other countries. And even though one ethnic minority group in the United States, Asian Americans, has significantly higher achievement scores than other minority groups, this group only represents about 4 percent of secondary school students, compared with about 32 percent of students who are black or Hispanic (as indicated in PISA 2006).

As noted, the paucity of ethnic status data makes it difficult to judge whether performance differences among US racial/ethnic groups are better, worse, or average compared with differences elsewhere. The evidence suggests that we could expect US gaps to be similar to gaps in other countries. However, assuming that US racial/ethnic gaps were significantly worse than in other countries would magnify the importance of race/ethnicity for the United States. Even if US racial/ethnic gaps were significantly narrower than those in other countries, the presence of almost any gap in combination with the substantially larger minority population in the United States makes those gaps far more consequential for overall US achievement. Any increase in US minority scores provides substantially greater benefit for overall US performance than similar increases in a typical developed country. For example, a 5 percent increase in the US minority score average on the PISA 2006 math test would result in a US average that is about 10 points higher (499 versus actual US average of 489). By contrast, assuming the same starting point for the average developed country (that is, an 80 scale-point gap between the majority average and any minority average) and a minority proportion of 8 percent, a 5 percent increase in minority score average would result in an average score that is only three points higher (511 versus actual developed average of 508).

Accounting for Racial and Ethnic Differences

Another objection to the argument that racial and ethnic gaps are important for understanding overall US performance in comparative perspective is the assertion that these gaps are driven by one or two other factors such as socioeconomic status (SES) or school quality. Because these factors are measured on

student and school questionnaires in the international assessments, one could argue that there is little need to consider race/ethnicity in assessing the United States' standing internationally. But achievement differences among US racial and ethnic groups do not simply reflect socioeconomic differences or other family background characteristics, or just differences in school quality. Rather, they reflect a combination of these factors along with other experiences in the broader environments (e.g., local neighborhoods) in which groups are likely to live. In addition, our understanding of even underlying family and school influences is greatly enhanced by also analyzing race and ethnicity, as a substantial body of research has shown.

First, SES and other background characteristics such as family structure usually account for some, but certainly not all, of the differences among racial and ethnic groups (Kao, Tienda, & Schneider, 1996; Kao & Thompson, 2003; Von Secker, 2004). The persistence of the black–white achievement gap, in particular, has been the subject of considerable research effort, and researchers routinely find that achievement differences remain even when comparing black and white students who have similar SES and family characteristics (Jencks & Phillips, 1998). Some research indicates that black–white differences at preschool ages can be explained by reference to family resources, but even this work shows that gaps at equivalent SES levels reemerge after only a few years of schooling (Fryer & Levitt, 2004).

Although other research suggests that differences between Hispanic and white students are accounted for by differences in SES (Warren, 1996), the higher performance of Asian students is only partially accounted for by the high educational levels of many Asian students' parents, and unexplained advantages remain (Sakamoto et al., 2010). In addition, differences in school or teacher practices do not, by themselves, explain racial and ethnic differences and, in fact, likely magnify early differences (Breen & Jonsson, 2005; Hanushek & Rivkin, 2009; Van de Werfhorst & Mijs, 2010). In reality, SES, family background, and school quality combine with additional elements such as heightened test anxiety, neighborhood safety and wealth, and friendship

networks to create encompassing educational environments that collectively provide fewer opportunities to disadvantaged groups (Brooks-Gunn, Duncan, Klebanov, & Sealand, 1993; Muller et al., 2010; Steele & Aronson, 1998).

Second, even if a set of family or school characteristics do account for gaps between specific groups (e.g., Hispanics and whites) or at specific times (e.g., preschool differences), the implications are not always straightforward. For example, SES is not a simple measure with a clear interpretation: it combines three distinct components—educational attainment, occupational status, and family income. In the international context, the latter is typically measured through reports of specific materials such as the number of books in the home (Buchmann, 2002). Although the three are strongly correlated, relatively high income levels and middle-level occupational statuses are not uncommon, particularly among working class families, and other small disjunctures among the three could mean that policies focusing on, say, income, neglect academic difficulties arising when parents have limited postsecondary education (such parents may have difficulty helping with their student's homework, high school course selection, and college applications) (Ferguson, 2007; Hill et al., 2004).

When SES does provide consistent internal measurement, it may still ignore important differences in the quality of attained degrees, the organizational context of a given job, and the accumulated wealth of families, all of which may contribute to the disadvantages black and Hispanic students experience compared with similar-SES whites and Asians (Braveman, Cubbin, Marchi, Egeter, & Chavez, 2001; Lareau, 1987, 2002).

Mandara, Varner, Greene, and Richman (2009), for example, found that a model of achievement that included measures of parents' academic orientation (such as enrichment activities and clear rules), as well as grandparents' education and SES, reduced the black–white gap among adolescents to zero. Similar arguments can be made about the meaning of other background factors such as the number of books in the home (what kinds of books?) and the nature of family structure (who else is involved in child-rearing?) (Brooks-Gunn, Klebanov, & Duncan,

1996; DeGarmo, Forgetch, & Martinez, 1999). These considerations are particularly important in the context of international assessments where limited space is available to gather detailed background data or explore subtle differences in school experiences.

Third, the exploration of race/ethnicity can contribute to a key debate about international achievement and student background: the varying SES–achievement relationship across countries. Analyses of international assessments show that achievement is more dependent on SES in the United States than in other countries—in other words, other countries do a better job of moderating the disadvantages that derive from limited family resources. To the extent that race/ethnicity is intertwined with SES and achievement in the United States to a greater degree than in other countries, understanding the SES–achievement relationship hinges on also understanding how the US SES–race/ethnicity relationship compares with that of other countries. More broadly, if researchers could compare differences in the association between family and school factors and achievement by ethnicity cross-nationally, they might be able to gauge what aspects of minority disadvantage are most amenable to policy intervention. This is, in fact, one of the central contributions of comparative research: the possibility of broadening our understanding of domestic processes by observing practices and outcomes in different environments (Rowan, 2002).

As a final comment, it is worth noting that (a) race/ethnicity is not an explanation in itself of achievement differences, unless one subscribes to discredited genetic explanations, and (b) racial/ethnic groups are not monolithic blocks but subsume an array of ethnic groups and groups with different immigration histories and nations of origin. Although race/ethnicity is not directly explained by other characteristics or experiences of students, a combination of other factors theoretically should, and occasionally does, explain specific achievement gaps. But the fact that there are different combinations of influences for individual racial/ethnic groups and these can further vary by age and school subject (not to mention other schooling outcomes) only strengthens the argument for a regular accounting of race and ethnicity in comparing US performance to the performance of other countries.

Implications and Research Recommendations

Boe and Shin (2005), Lowell and Salzman (2007), and other authors (Bracey, 1996, 2005; Westbury, 1993) have noted that the diversity of the US population creates challenges for US education that many other countries do not have to face. Indeed, as we have shown, the United States possesses a substantially more diverse student population than other countries, with large gaps between high-achieving and low-achieving groups. The US gap, in particular, may be bound to a host of other conditions and experiences that impair learning and are not subject to easy solutions. However, these authors tend to view racial and ethnic differences as explaining away overall US performance on international assessments, despite the authors' desire to eliminate the differences. Boe and Shin, for example, although noting that both nonschool (family) and school system factors may contribute to racial/ethnic differences, subsequently state that the aspiration for the United States to be "first in the industrialized world" is "unreasonable." They write:

The US is not "first in the industrialized world" in minimizing the percentage of its population living in poverty or in minimizing its infant mortality rate. So why should anyone expect the US to be first in the world in educational achievement? There is, after all, abundant evidence that these types of social indicators are strongly associated with educational achievement. (p. 694)

Boe and Shin do not state what they believe the US goal should be, though they do make a generic statement that "there is always room for improvement" (p. 694). The unfortunate implication is that there is little that school systems can do about the social factors that disadvantage some groups of students. This perspective misses entirely the fact that certain groups in the United States do, in fact, perform at levels that regularly make them among the first in the world, as the results here show.

Likewise, despite acknowledging the undesirability of racial/ethnic gaps and recommending a general policy of improving achievement among low-performing groups, Lowell and Salzman (2007) write

that "[o]ne could argue that it is the diversity and openness of the United States that both contribute to its high economic performance and its lower average educational performance" (p. 24). Yet this statement also misses the fact that US diversity coexists with high average educational performance among minority groups such as Asian Americans (as well as high achievement among subgroups of Hispanics and blacks). A minority background is clearly not incompatible with both high achievement and economic success.

Therefore, the significant gaps among racial/ethnic groups shown on international assessments are a cause of concern more than relief. Even under the recognition that overall US performance is not poor, there is a long way between "anything less than first place is regarded as poor performance" (Boe & Shin, 2005, p. 694) and a dissatisfaction with overall middling achievement.

More importantly, recognizing that US performance cannot be adequately characterized by reference to overall standings and averages—that there is a sharp and dismaying gap in how US racial and ethnic groups perform—leads to a conclusion often made by others: that low-scoring minority groups should be the target of educational improvement efforts. Doing so would be more efficient and more effective at improving the United States' standing on international achievement than retooling entire curricula, school organizational structures, teacher training regimens, or other complete transformations of the educational system (e.g., Stedman, 1997) that are based on the false assumption that US achievement is uniformly mediocre.

However, though the comparative results presented here show distinctly separate patterns of achievement for different racial and ethnic groups, the evidence from international assessments is too broad to support clear policy prescriptions, as others have noted (Cavanagh & Manzo, 2009; Smith, 2002). Therefore, using international assessment results to effectively inform reform goals and policy proposals for low-scoring minority groups requires improving the comparative research base first. Here, we present two research recommendations that would

further understanding of the US racial and ethnic divide and lead to innovative policy proposals. The first recommendation addresses limits to the international datasets themselves, while the second recommendation focuses on research directions that can be pursued with the current data.

1. The IEA and OECD should make greater efforts to collect and provide assessment data by minority ethnic status.

This recommendation is the most immediate and direct to arise from this discussion. The IEA and OECD should, with the encouragement of the United States, attempt to gather data reflecting key ethnic or cultural groupings for a variety of reasons. As noted, US research and results from other countries indicate a strong influence of ethnic background on educational achievement (Jencks & Phillips, 1998; Kalmijn & Kraaykamp, 2003; Kao & Thompson, 2003; Shavit, 1990). This influence is not solely explained by or consistent with the small set of additional variables (such as SES, immigration status, or language use) included on international assessment background surveys.

In addition, ethnic population sizes vary substantially by country, and the relationship between ethnicity and achievement may vary across countries. Appropriate ethnic indicators would identify major groups based on cultural identities rooted in religion, historical experience, country of origin, and/or language. It would derive from countries' own classification systems and not impose other nations' categories on individual countries. Appropriately defined and documented, this information could serve as a valuable resource for countries to compare the progress and status of their minority groups.

Such an effort is both practically and politically difficult, however. As NCES (2010) itself notes, "certain demographic characteristics are not easy to collect across countries. Race/ethnicity is one of these. Moreover, even if the data were collected by all countries, comparisons may not be meaningful because the makeup of minority populations in each country differ[s]." But a number of constructs are difficult to measure cross-nationally. The measure of SES used in the international assessments, for

example, allows countries to adapt lists of household items (to help capture differences in income) to their societal context. Assessment scores themselves are very difficult to measure cross-nationally, but assessment items are carefully designed not to introduce bias or cause confusion for different national populations. Difficulty is not a primary stumbling block in measuring race/ethnicity. Comparability is a larger concern, yet data such as immigrant status are routinely collected despite similar issues in the size and makeup of immigrant populations across countries.

Further, although it may not be easy to measure minority status in comparable fashion, it would be easy for each country to measure minority membership in its own manner. These data could then be added as part of the international databases. Currently, minority status data collected by the United States, Australia, New Zealand, and possibly others are not even included in the combined international datasets. At the least, IEA and OECD could add these existing variables and thereby support initial efforts at understanding cross-national differences in the relationship between minority status and achievement.

A larger issue is that identifying ethnic or minority status may be politically difficult. Some countries, such as France, notably do not collect, and have policies against collecting, racial or ethnic data. A similar fear in other countries may be that uncovering systematic disadvantages among specific populations might result in special programs and favors that undermine cultural unity and provoke sectarian discord. However, all participating countries have engaged in gathering socioeconomic, immigration, and language-use data that have been used to suggest precisely this kind of targeted educational support.

In addition, ignoring systematic disadvantages among certain clearly identifiable groups raises the risk of reinforcing nonmajoritarian and extra-national identities that undermine the cultural unity supposedly achieved by maintaining formal recognition of only broad categories of citizenship or national membership. Nonetheless, it must be noted that educational assessment programs are unlikely

levers through which data collection policies such as these will be changed. However, international effort could be devoted to adding an ethnic indicator among a subset of interested developed and developing countries. Country-specific measures from even a limited number of countries would be extremely valuable to researchers in the United States and abroad.

2. Comparative and domestic researchers should expand their examination of race/ethnicity in the United States across multiple administrations of the international assessments and into specific subdomains of achievement.

Because of the lack of comparable ethnic or minority status data (other than that on immigrant students), US data from TIMSS, PIRLS, and PISA have rarely been used to explore racial and ethnic differences in achievement and possible reasons for those differences. The richness of domestic sources of data, which may involve more in-depth background data, longitudinal or experimental designs, and greater representation of minority groups, makes the use of the cross-sectional, broad-based international assessments less than ideal for investigating the causal relationship between race/ethnicity and achievement. However, the US assessment datasets can contribute to research in specific ways. Researchers can use the accumulated evidence across multiple administrations to explore racial/ethnic patterns at given ages or grades, patterns for specific subjects, trends in achievement among racial/ethnic groups, or changes in racial/ethnic effects over time. Though US national datasets such as the National Assessment for Educational Progress (NAEP) allow these kinds of comparisons, the international assessments provide additional data, cover different assessment concepts, and enable benchmarking results to other countries. The international assessments also survey an additional age (15), which none of the NAEP programs target.

Within the international assessments' content foci, researchers can take advantage of one assessment's major strengths: detailed performance data (including item-level responses) in specific domains of math, science, reading, and other subjects. Little is known about whether groups have certain strengths or

weaknesses in individual areas or with specific types of material from the international assessments. NCES, for example, does not typically publish detailed information about the distribution of international achievement scores by race and ethnicity or the percentage of different groups achieving specific international assessment benchmarks (a subsidiary recommendation here is for NCES itself to publish more information about US students' performance by race and ethnicity on international tests). Identifying areas of strength and weakness may help foster targeted interventions and reconsideration of how specific topics are taught to different US audiences, as well as what international curriculum or instructional models may be successful with US minorities.

Further, more could be done to explicitly compare racial/ethnic effects across countries. In addition to the United States, New Zealand has multiple years of racial/ethnic data, Australia has data on their indigenous population, and other countries may have existing data to use for exploration of ethnic differences (for example, for PISA 2000, the Netherlands reported immigrant data by specific national origins—see Wijnstra, 2001).

Creative uses of immigrant status, language use, and geographic locale might be discovered that can shed light on processes of ethnic achievement cross-culturally. Understanding these processes from a broader perspective can help US researchers identify aspects of the organization of US schools or classrooms that are not apparent from a solely national perspective. For example, do countries with strong differentiation among secondary school programs have minority populations that achieve at higher or lower levels than minority groups in the United States, and what implications would that have for the role of career and technical education in the United States? How does US minority performance compare with minority performance in countries with comprehensive final exit exams or other school policies? The variety of national systems and practices encourages policy makers and researchers to think beyond the range of variation available within the United States.

Conclusion

In their insightful chapter about analyzing international assessments, Raudenbush and Kim (2002) note that “mean differences will be misleading when statistical interactions are present. . . . Two nations that look similar on average, for example, could differ dramatically if subgroups were compared” (p. 271). The analysis I present here is striking evidence for that observation. US achievement on international assessments is sharply divided between a group of high-achieving Asian and white students and a group of low-achieving black and Hispanic students. The lowest-performing racial and ethnic groups perform at levels below almost every other developed country, while the highest-performing groups are often at or near the top of international achievement rankings. This divided pattern is particularly important for the United States because of its relatively larger minority population

and the lack of simple explanations for racial and ethnic disparities.

In short, the United States is neither a poor, average, nor above-average country when it comes to the educational performance of its youth. Rather, it is a country characterized by world-class achievement by some of its students and dismaying performance by a notable minority. Our divided performance makes overarching assertions about general US achievement unhelpful from an analytical or policy perspective. With additional research comparing the performance of US racial/ethnic groups with similar groups in other developed countries, identifying specific areas of weakness within individual groups, and linking local patterns to national ones, the United States has the potential to build a stronger foundation for effective educational policy and practice than is currently being derived from international assessment programs.

References

- Baer, J., Baldi, S., Ayotte, K., & Green, P. J. (2007). *The reading literacy of US fourth-grade students in an international context results from the 2001 and 2006 Progress in International Reading Literacy Study (PIRLS)*. (NCES 2008-017). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, US Department of Education.
- Baldi, S., Ying Jin, Y., Skinner, M., Green, P. J., & Herget, D. (2007). *Highlights from PISA 2006: Performance of US 15-year-olds students in science and mathematics literacy in an international context* (NCES 2008-016). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, US Department of Education.
- The Bill & Melinda Gates Foundation. (2011). Top 10 facts about college readiness. Retrieved March 10, 2011, from <http://www.gatesfoundation.org/college-ready-education/Pages/10-fast-facts-college-ready-education.aspx>
- Boe, E. L., & Shin, S. (2005). Is the United States really losing the international horse race in academic achievement? *Phi Delta Kappan*, 86(9), 688–695.
- Bracey, G. W. (1996). International comparisons and the condition of American education. *Educational Researcher*, 25(1), 5–11.
- Bracey, G. W. (2005). Plus ça change. *Phi Delta Kappan*, 86(6), 476–478.
- Braveman, P., Cubbin, C., Marchi, K., Egeter, S., & Chavez, G. (2001). Measuring socioeconomic status/position in studies of racial/ethnic disparities: Maternal and infant health. *Public Health Reports*, 116(5), 449–463.
- Breen, R., & Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology*, 31, 223–243.
- Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K., & Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology*, 99(2), 353–395.

- Brooks-Gunn, J., Klebanov, P. K., & Duncan, G. J. (1996). Ethnic differences in children's intelligence scores: Role of economic deprivation, home environment, and maternal characteristics. *Child Development, 67*, 396–408.
- Buchmann, C. (2002). Measuring family background in international studies of education: Conceptual issues and methodological challenges. In A. C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 150–197). Washington, DC: National Academy Press.
- Bussière, P., Knighton, T., & Pennock, D. (2007). *Measuring up: The Canadian results of the OECD PISA study. The performance of Canada's youth in science, reading, and mathematics: 2006 first results for Canadians aged 15*. Ottawa, Canada: Statistics Canada.
- Cavanagh, S., & Manzo, K. K. (2009). International exams yield less-than-clear lessons. *Education Week, 28*(29), 16–17.
- Caygill, R. (2008). *Trends in year 5 science achievement 1994 to 2006: New Zealand results from three cycles of the Trends in International Mathematics and Science Study (TIMSS)*. Wellington, New Zealand: Research Division, Ministry of Education.
- Caygill, R., & Kirkham, S. (2008). *Trends in year 5 mathematics achievement 1994 to 2006: New Zealand results from three cycles of the Trends in International Mathematics and Science Study (TIMSS)*. Wellington, New Zealand: Research Division, Ministry of Education.
- Central Intelligence Agency. (2010). *The world factbook 2010*. Retrieved February 10, 2011, <https://www.cia.gov/library/publications/the-world-factbook/index.html>
- Council on Competitiveness. (2005). *Innovate America: Thriving in a world of challenge and change*. Washington, DC: Council on Competitiveness.
- Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. New York: Teachers College Press.
- DeGarmo, D. S., Forgatch, M. S., & Martinez, Jr., C. R. (1999). Parenting of divorced mothers as a link between social status and boys' academic outcomes: Unpacking the effects of socioeconomic status. *Child Development, 70*, 1231–45.
- Editorial: Our view on education: "We're no. 15!" Doesn't cut it in today's global economy. [Editorial]. (2010, December 15). *USA Today*.
- Ferguson, R. (2007). Parenting practices, teenage lifestyles, and academic achievement among African American children. *Focus, 25*(1), 18–26.
- Fryer, Jr., R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics, 86*(2), 447–464.
- Gonzales, P., Guzman, J. C., Partelow, P., Pahlke, E., Jocelyn, L., Kastberg, D., & Williams, T. (2004). *Highlights from the Trends in International Mathematics and Science Study (TIMSS) 2003* (NCES 2005-005). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, US Department of Education.
- Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of US fourth- and eighth-grade students in an international context* (NCES 2009-001). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, US Department of Education.
- Hanushek, E. A., Peterson, P. E., & Woessmann, L. (2010). Teaching math to the talented. *Education Next, 11*(1). Retrieved January 19, 2011, from http://educationnext.org/files/ednext_20111_TeachingMathToTheTalented.pdf
- Hanushek, E. A., & Rivkin, S. G. (2009). Harming the best: How schools affect the black-white achievement gap. *Journal of Policy Analysis and Management, 28*, 366–696.

- Hill, N. E., Castellino, D. R., Lansford, J. E., Nowlin, P., Dodge, K. A., Bates, J. E., & Pettit, G. S. (2004). Parent academic involvement as related to school behavior, achievement, and aspirations: Demographic variations across adolescence. *Child Development, 75*(5), 1491–1509.
- Huang, M.-H. (2009). Beyond horse race comparisons of national performance averages: Math performance variation within and between classrooms in 38 countries. *Educational Research and Evaluation, 15*(4), 327–342.
- International Monetary Fund. (2010). World economic outlook: *Recovery, risk, and rebalancing*. Washington, DC: International Monetary Fund.
- Jencks, C., & Phillips, M., Eds. (1998). *The black-white test score gap*. Washington, DC: Brookings Institution Press.
- Kalmijn, M., & Kraaykamp, G. (2003). Dropout and downward mobility in the educational career: An event-history analysis of ethnic schooling differences in the Netherlands. *Educational Research and Evaluation, 9*, 265–287.
- Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology, 29*, 417–442.
- Kao, G., Tienda, M., & Schneider, B. (1996). Racial and ethnic variation in educational achievement. *Research in Sociology of Education and Socialization, 11*, 263–297.
- Koretz, D. (2009). How do American students measure up? Making sense of international comparisons. *Future of Children, 19*(1), 37–51.
- Lareau, A. (1987). Social class differences in family-school relationships: The importance of cultural capital. *Sociology of Education, 60*, 73–85.
- Lareau, A. (2002). Invisible inequality: Social class and childrearing in black and white families. *American Sociological Review, 67*, 747–776.
- Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Williams, T., Kastberg, D., & Jocelyn, L. (2004). *International outcomes of learning in mathematics literacy and problem solving: PISA 2003 Results from the US perspective* (NCES 2005-003). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, US Department of Education.
- Lowell, B. L., & Salzman, H. (2007). *Into the eye of the storm: Assessing the evidence on science and engineering education, quality, and workforce demand*. Washington, DC: The Urban Institute.
- Mandara, J., Varner, F., Greene, N., & Richman, S. (2009). Intergenerational predictors of the black-white achievement gap. *Journal of Educational Psychology, 101*(4), 867–878.
- Muller, C., Riegle-Crumb, C., Schiller, K. S., Wilkinson, L., & Frank, K. A. (2010). Race and academic achievement in racially diverse high schools: Opportunity and stratification. *Teachers College Record, 112*(4), 4–5.
- National Academy of Science. (2007). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: Committee on Science, Engineering, and Public Policy (COSEPUP).
- National Center for Educational Statistics. (2010). Frequently asked questions about the assessment. Washington, DC: National Center for Educational Statistics. Retrieved February 9, 2011, from <http://nces.ed.gov/timss/faq.asp>
- Ogle, L. T., Sen, A., Pahlke, E., Jocelyn, L., Kastberg, D., Roey, S., & Williams, T. (2003). *International comparisons in fourth-grade reading literacy: Findings from the Progress in International Reading Literacy Study (PIRLS) of 2001* (NCES 2003-073). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, US Department of Education.
- Organisation for Economic Co-operation and Development (OECD). (2007). *PISA 2006: Science competencies for tomorrow's world: Vol. 1, analysis*. Paris: OECD.

- Raudenbush, S. W., & Kim, J.-S. (2002). Statistical issues in analysis of international comparisons of educational achievement. In A. C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 267–294). Washington, DC: National Academy Press.
- Rowan, B. (2002). Large-scale, cross-national surveys of educational achievement: Promises, pitfalls, and possibilities. In A. C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 321–351). Washington, DC: National Academy Press.
- Sakamoto, A., Goyette, K. A., & Kim, C. (2010). Socioeconomic attainments of Asian Americans. *Annual Review of Sociology*, 35, 255–276.
- Shavit, Y. (1990). Segregation, tracking, and the educational attainment of minorities: Arabs and Oriental Jews in Israel. *American Sociological Review*, 55, 115–126.
- Smith, M. S. (2002). Drawing inferences for national policy from large-scale cross-national education surveys. In A. C. Porter and A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 295–317). Washington, DC: National Academy Press.
- Stedman, L. C. (1997). International achievement differences: An assessment of a new perspective. *Educational Researcher*, 26(3), 4–15.
- Steele, C., & Aronson, J. (1998). Stereotype threat and the test performance of academically successful African Americans. In C. Jencks & M. Phillips (Eds.), *The black-white test score gap* (pp. 401–430). Washington, DC: The Brookings Institute.
- Thomsen, S., McKelvie, P., & Murnane, H. (2006). *Achievement of Australia's early secondary indigenous students: Findings from TIMSS 2003*. Camberwell, Australia: Australian Council for Educational Research.
- Thomsen, S., Wernert, N., Underwood, C., & Nicholas, M. (2008). *TIMSS 07: Taking a closer look at mathematics and science in Australia*. Camberwell, Australia: Australian Council for Educational Research.
- United Nations Educational, Scientific, and Cultural Organization (UNESCO). (2010). *Global education digest 2010: Comparing education statistics around the globe*. New York: UNESCO.
- Van de Werfhorst, H. G., & Mijs, J. J. B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, 36, 407–428.
- Von Secker, C. (2004). Science achievement in social contexts: Analysis from the National Assessment of Educational Progress. *The Journal of Educational Research*, 98(2), 67–78.
- Warren, J. R. (1996). Educational inequality among white and Mexican-origin adolescents in the American Southwest: 1990. *Sociology of Education*, 69, 142–158.
- Westbury, I. (1993). American and Japanese achievement...again: A response to Baker. *Educational Researcher*, 22(3), 21–25.
- Wijnstra, J. M. (2001). *Bruikbare kennis en vaardigheden voor jonge mensen*. Arnhem, Netherlands: Citogroep.

Acknowledgments

The author would like to thank Patricia J. Green and the RTI Press reviewers for their helpful comments on this paper.

RTI International is an independent, nonprofit research organization dedicated to improving the human condition by turning knowledge into practice. RTI offers innovative research and technical solutions to governments and businesses worldwide in the areas of health and pharmaceuticals, education and training, surveys and statistics, advanced technology, international development, economic and social policy, energy and the environment, and laboratory and chemistry services.

The RTI Press complements traditional publication outlets by providing another way for RTI researchers to disseminate the knowledge they generate. This PDF document is offered as a public service of RTI International.