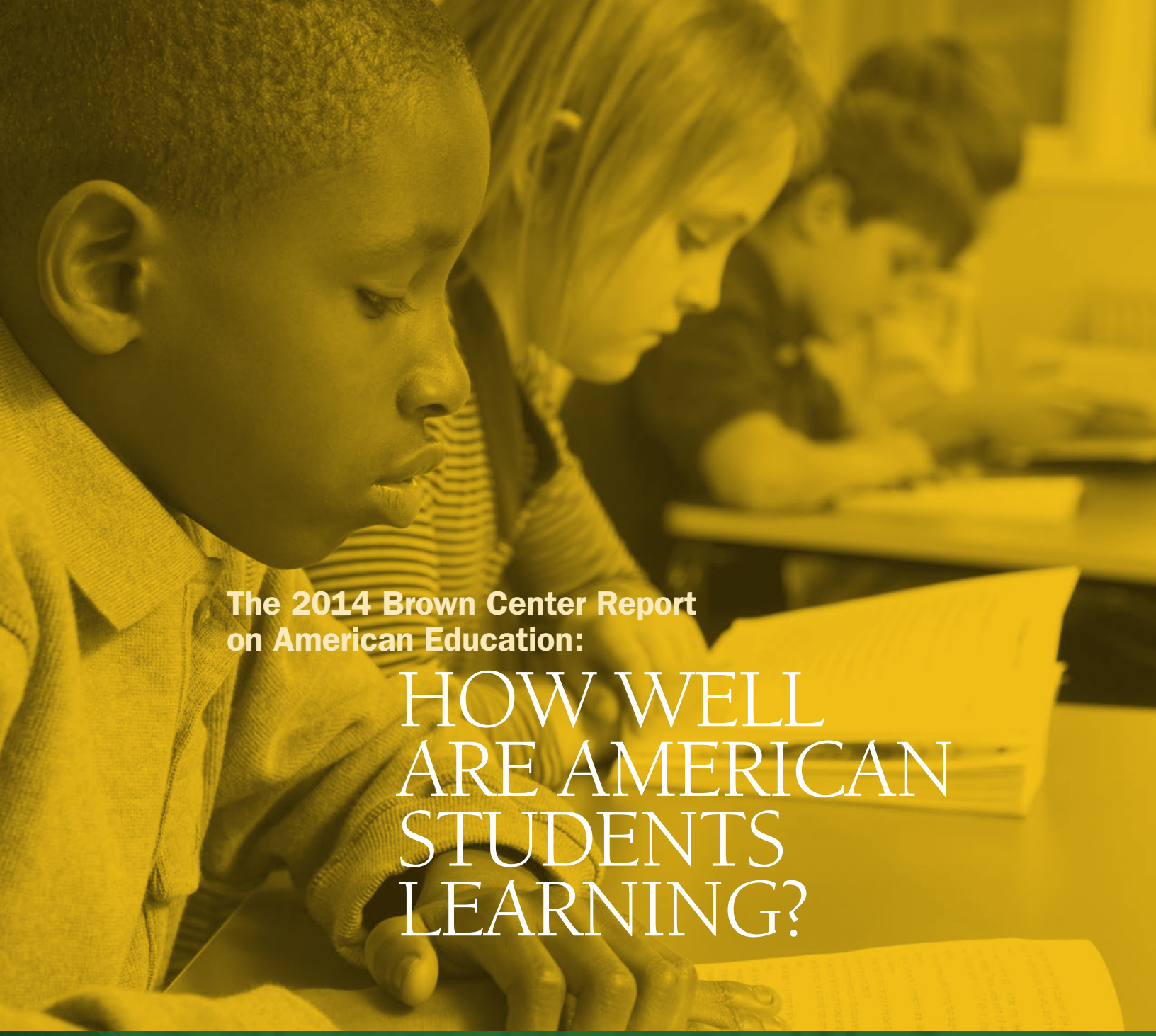


The 2014 Brown Center Report on American Education:

HOW WELL ARE AMERICAN STUDENTS LEARNING?

With sections on the PISA-Shanghai Controversy, Homework, and the Common Core

BROWN CENTER on Education Policy at BROOKINGS

ABOUT BROOKINGS

The Brookings Institution is a private nonprofit organization devoted to independent research and innovative policy solutions. For more than 90 years, Brookings has analyzed current and emerging issues and produced new ideas that matter—for the nation and the world.

ABOUT THE BROWN CENTER ON EDUCATION POLICY

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

This report was made possible by the generous financial support of The Brown Foundation, Inc., Houston.

The 2014 Brown Center Report on American Education:

HOW WELL ARE AMERICAN STUDENTS LEARNING?

With sections on the PISA-Shanghai Controversy, Homework, and the Common Core

March 2014
Volume 3, Number 3

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

TABLE OF CONTENTS

Research assistance by: Mike Gallaher,
Katharine Lindquist, and Sarah Whitfield

THE 2014 BROWN CENTER REPORT ON AMERICAN EDUCATION

This year's Brown Center Report on American Education represents the third installment of volume three and the 13th issue overall since the publication began in 2000. Three studies are presented. All three revisit a topic that has been investigated in a previous Brown Center Report. The topics warrant attention again because they are back in the public spotlight.

Part one summarizes the recent controversy involving the Programme for International Student Assessment (PISA) and its treatment of Shanghai-China. The PISA is a test given to 15-year-olds every three years in math, reading, and science. Sixty-five national and subnational jurisdictions participated in the 2012 PISA. When the scores were released in December 2013, no one was surprised that Shanghai-China scored at the top in all subjects. But what has been overlooked by most observers—and completely ignored by the authorities running PISA—is that Shanghai's population of 15-year-olds is sifted and shaped in ways that make its scores incomparable to those of any other participant.

China requires all citizens to hold a *hukou*, a passport-like document issued by a family's province of origin. The system dates back to 1958 and the authoritarian regime of Mao Zedong. The original purpose of hukou was to control where people lived. Today it serves the purpose of rationing social services, including health care and education. Large cities in China are inundated with migrants who leave poor, rural areas in search of work. Admission to an academic high school in Shanghai is almost impossible for a student not holding a Shanghai hukou. In addition, students can only take the *gaokao*, the national college entrance exam, in their province of hukou registration. As a consequence, tens of thousands of Shanghai families send their children back to rural villages as the children approach high

school age. The only other option is to leave the children behind in the first place, the fate of approximately 60 million children nationwide.

Hukou is hereditary. Children born in Shanghai to migrant parents are not entitled to a Shanghai hukou. In 2012, Zhang Haite, a 15-year-old student in Shanghai, took to the internet to protest being sent away to a rural village for high school, despite the fact that she had never lived there. The hukou system has been condemned by Human Rights Watch and Amnesty International for its cruelty in breaking up families and for limiting the educational opportunities of children based on their family's hukou status. Not only has PISA been silent on the impact of hukou on the composition of Shanghai's 15-year-old population, but PISA documents have also repeatedly held up Shanghai as a model of educational equity and praised its treatment of disadvantaged children.

From October 2013 to January 2014, a series of three essays on the Brown Center Chalkboard criticized PISA for ignoring the devastating effects of the hukou system. PISA officials were also criticized for several contradictory statements that cloak China's participation in PISA in a cloud of secrecy. PISA officials and defenders of PISA responded to the critique. Part one summarizes the debate and offers lessons that the affair offers for PISA's future governance. Several steps need to be taken to restore PISA's integrity.

Part two is on homework, updating a study presented in the 2003 Brown Center Report. That study was conducted at a time when homework was on the covers of several popular magazines. The charge then was that the typical student's homework load was getting out of control. The 2003 study examined the best evidence on students' homework burden and found the charge to be an exaggeration.

Now, a little more than a decade later, homework is again under attack. In 2011, the *New York Times* ran a front page story describing "a wave of

districts across the nation trying to remake homework amid concerns that high stakes testing and competition for college have fueled a nightly grind that is stressing out children and depriving them of play and rest, yet doing little to raise achievement, especially in elementary grades."[1] A September 2013 *Atlantic* article, "My Daughter's Homework is Killing Me," featured a father who spent a week doing the same three or more hours of nightly homework as his daughter.

The current study finds little evidence that the homework load has increased for the average student. Those with a heavy burden, two or more hours of homework per night, do indeed exist, but they are a distinct minority. The maximum size of the heavy homework group is less than 15%, and that's true even for 17-year-olds. In national polls, parents are more likely to say their children have too little homework than too much. And a solid majority says the amount of their children's homework is about right. With one exception, the homework load has remained stable since 1984. The exception involves 9-year-olds, primarily because the percentage of 9-year-olds with no homework declined while the percentage with some homework—but less than an hour—increased.

Part three is on the Common Core State Standards (CCSS). Forty-five states have signed on to the Common Core and are busy implementing the standards. How is it going? Admittedly, the Common Core era is only in the early stages—new tests and accountability systems based on the standards are a couple of years away—but states have had three or four years under the standards. Sufficient time has elapsed to offer an early progress report.

The progress report proceeds along two lines of inquiry. First, a ranking system crafted by researchers at Michigan State University is employed to evaluate progress on NAEP from 2009–2013. The MSU experts found that states with math standards that were similar to the Common Core in 2009 scored higher on the eighth grade NAEP that year compared to states

with standards dislike the Common Core. The current study examines data from the NAEP tests conducted in 2011 and 2013 and asks whether the same finding holds for subsequent changes in NAEP scores. Have the states with CCSS-like standards made greater gains on the eighth grade NAEP since 2009? It turns out they have not.

The second line of inquiry utilizes a rubric that categorizes each state on the strength of its implementation of CCSS. NAEP gains were again compared. Here the news is more encouraging for the Common Core. States with stronger implementation of the CCSS have made larger NAEP gains. The downside to this optimistic finding is that the difference is quite small. If Common Core is eventually going to fulfill the soaring expectations of its supporters, much greater progress must become evident.

Part I LESSONS FROM THE PISA-SHANGHAI CONTROVERSY

THE RESULTS OF THE 2012 PROGRAMME FOR INTERNATIONAL Student Assessment (PISA) were released in December 2013. The test is administered to 15-year-olds every three years by a division of the Organisation for Economic Co-operation and Development (OECD) in Paris. Scores were reported in math, reading, and science for more than sixty nations and subnational jurisdictions. The top scoring participant, as in 2009, was Shanghai-China in all three subjects. A controversy erupted concerning Shanghai's participation in PISA. A series of Brown Center Chalkboard essays took part in the debate.[2]

This section of the Brown Center Report will discuss the lessons that can be learned from the controversy. International tests are an increasingly important source of information on the performance of school systems. The response of PISA officials to questions raised about Shanghai highlights serious flaws in the governance of PISA that should be addressed.

Let's start with a brief summary of the controversy.

Hukou and Shanghai's Missing 15-Year-Olds

The Chinese hukou system is unique in the world. Started by Mao Zedong in 1958 as a tool for controlling internal migration from rural to urban areas, hukou is a household registration system that restricts rural migrants' access to urban social services, including education. Age 15 is a pivotal year in the life of Chinese adolescents. Compulsory education ends at the end of ninth grade and students must take the *zhongkao*, the senior high school entrance exam, to determine their high school of attendance. Thanks to recent reforms, migrants without a Shanghai hukou can now enroll children in public primary and middle schools, but admission to academic high schools is severely restricted. In addition, students can only sit for the *gaokao*, the national college entrance exam, in the province of hukou registration. As a consequence, tens of thousands of families send their children back to rural villages as the children approach high school age.[3]

The other option is to leave children behind with relatives when parents relocate to cities in search of work. Approximately 60 million children in China are "left-behinds."[4] An emerging literature in China documents the psychological damage done to both children and parents as families are forced apart.

Hukou is hereditary. Second generation migrants, children who are born in cities to migrant parents, are not granted an urban hukou. In 2012, a 15-year-old Shanghai girl, Zhan Haite, organized an internet-based campaign to protest being forced to attend high school in Jiangxi province.[5] She was locked out of Shanghai's high schools. Although she had attended both primary and middle school in Shanghai, and although she was born in Guangdong, another large city, her family's ancestral home is rural Jiangxi. Her hukou is from Jiangxi province.

Hukou acts as a giant sifting machine, barring or driving out migrant children from urban schools. Coincidentally, PISA assesses the academic achievement of 15-year-olds. At the same time migrant families in Shanghai are feeling the brunt of the hukou system, a random sample is drawn from children attending schools in Shanghai for the PISA test. There is nothing wrong with PISA's sampling technique, but even a pristine sampling strategy cannot compensate for a population that has already been culled of migrant students.

The impact of hukou on the 15-year-old population in Shanghai is easily seen in PISA data. Table 1-1 shows all of the participants in the 2012 PISA. On average, 15-year-olds constitute 1.2872% of the participating nations' total population. When this average is used to predict the number of 15-year-olds in each jurisdiction—and then the countries are ranked by the deviation of the reported number of 15-year-olds to the predicted number—two outliers stand out at the bottom of the table. The United Arab Emirates' outlier status can be explained because less than 15% of the UAE's population are citizens—the rest are from foreign

Analysis of 15-Year-Old Populations in Participating PISA Countries

Table 1-1

	National Population (mil)	15-Year-Olds Reported by PISA	15-Year-Olds as % of POP	Predicted Number of 15-year-olds	Difference	PISA's Reported 15-year-olds as % of Predicted
Albania	3.2	76,910	2.40344	41,191	35,719	186.72%
Jordan	6.1	129,492	2.12282	78,520	50,972	164.92%
Peru	29.5	584,294	1.98066	379,729	204,565	153.87%
Vietnam	86.9	1,717,996	1.97698	1,118,593	599,403	153.59%
Colombia	45.5	889,729	1.95545	585,684	304,045	151.91%
Malaysia	28.3	544,302	1.92333	364,283	180,019	149.42%
Mexico	112.3	2,114,745	1.88312	1,445,546	669,199	146.29%
Brazil	193.3	3,574,928	1.84942	2,488,193	1,086,735	143.68%
Costa Rica	4.6	81,489	1.77150	59,212	22,277	137.62%
Indonesia	237.6	4,174,217	1.75683	3,058,431	1,115,786	136.48%
Turkey	73.7	1,266,638	1.71864	948,680	317,958	133.52%
Argentina	40.5	684,879	1.69106	521,323	163,556	131.37%
Chile	17.1	274,803	1.60704	220,114	54,689	124.85%
Uruguay	3.4	54,638	1.60700	43,765	10,873	124.84%
Kazakhstan	16.4	258,716	1.57754	211,104	47,612	122.55%
Israel	7.6	118,953	1.56517	97,829	21,124	121.59%
Thailand	63.9	982,080	1.53690	822,532	159,548	119.40%
Iceland	0.3	4,505	1.50167	3,862	643	116.66%
Montenegro	0.6	8,600	1.43333	7,723	877	111.35%
Chinese Taipei	23.2	328,356	1.41533	298,635	29,721	109.95%
Korea	49.4	687,104	1.39090	635,886	51,218	108.05%
New Zealand	4.4	60,940	1.38500	56,638	4,302	107.60%
Norway	4.9	64,917	1.32484	63,074	1,843	102.92%
Ireland	4.5	59,296	1.31769	57,925	1,371	102.37%
Denmark	5.5	72,310	1.31473	70,797	1,513	102.14%
Australia	22.3	291,967	1.30927	287,050	4,917	101.71%
United States	309.3	3,985,714	1.28862	3,981,366	4,348	100.11%
France	62.8	792,983	1.26271	808,373	−15,390	98.10%
Tunisia	10.5	132,313	1.26012	135,158	−2,845	97.90%
Luxembourg	0.5	6,187	1.23740	6,436	−249	96.13%
Canada	34.1	417,873	1.22543	438,941	−21,068	95.20%
Hong Kong–China	7.0	84,200	1.20286	90,105	−5,905	93.45%
United Kingdom	62.0	738,066	1.19043	798,075	−60,009	92.48%

continued

Table 1-1 (cont.)

continued

	National Population (mil)	15-Year-Olds Reported by PISA	15-Year-Olds as % of POP	Predicted Number of 15-year-olds	Difference	PISA's Reported 15-year-olds as % of Predicted
Netherlands	16.6	194,000	1.16867	213,678	−19,678	90.79%
Lithuania	3.3	38,524	1.16739	42,478	−3,954	90.69%
Finland	5.4	62,523	1.15783	69,510	−6,987	89.95%
Belgium	10.8	123,469	1.14323	139,020	−15,551	88.81%
Switzerland	7.8	87,200	1.11795	100,403	−13,203	86.85%
Hungary	10.0	111,761	1.11761	128,722	−16,961	86.82%
Poland	38.2	425,597	1.11413	491,717	−66,120	86.55%
Austria	8.4	93,537	1.11354	108,126	−14,589	86.51%
Slovak Republic	5.4	59,723	1.10598	69,510	−9,787	85.92%
Macao-China	0.6	6,600	1.10000	7,723	−1,123	85.46%
Sweden	9.3	102,087	1.09771	119,711	−17,624	85.28%
Serbia	7.3	80,089	1.09711	93,967	−13,878	85.23%
Croatia	4.4	48,155	1.09443	56,638	−8,483	85.02%
Singapore	5.1	53,637	1.05171	65,648	−12,011	81.70%
Portugal	10.6	108,728	1.02574	136,445	−27,717	79.69%
Italy	60.3	605,490	1.00413	776,193	−170,703	78.01%
Greece	11.3	110,521	0.97806	145,456	−34,935	75.98%
Germany	81.8	798,136	0.97572	1,052,945	−254,809	75.80%
Slovenia	2.0	19,471	0.97355	25,744	−6,273	75.63%
Estonia	1.3	12,649	0.97300	16,734	−4,085	75.59%
Japan	128.1	1,241,786	0.96939	1,648,927	−407,141	75.31%
Bulgaria	7.6	70,188	0.92353	97,829	−27,641	71.75%
Czech Republic	10.5	96,946	0.92330	135,158	−38,212	71.73%
Spain	46.0	423,444	0.92053	592,120	−168,676	71.51%
Cyprus	1.1	9,956	0.90509	14,159	−4,203	70.31%
Russian Federation	141.9	1,272,632	0.89685	1,826,563	−553,931	69.67%
Latvia	2.2	18,789	0.85405	28,319	−9,530	66.35%
Qatar	1.7	11,667	0.68629	21,883	−10,216	53.32%
Romania	21.5	146,243	0.68020	276,752	−130,509	52.84%
United Arab Emirates	8.3	48,824	0.58824	106,839	−58,015	45.70%
Shanghai–China	23.0	108,056	0.46981	296,060	−188,004	36.50%
AVERAGE	35.4	497,728	1.28722			

Source: Author's calculations

nations.[6] But notice who sits below the UAE and at the bottom of the chart—Shanghai-China. Using the 1.2872% international average as a guide, one would expect about 300,000 15-year-olds in Shanghai, a province of 23 million people. Instead, only about one-third of that amount, 108,056, is reported by PISA.

As mentioned above, migrants either leave their children behind in rural villages or bring them to Shanghai knowing that they will have to be sent away sometime before the children attend high school. The left behind population specifically related to Shanghai is difficult to pinpoint, but noted hukou expert Kam Wing Chan offers 200,000 to 300,000 as a reasonable figure.[7] The steadily declining population of migrant children in Shanghai as they grow older can be seen in Figure 1-1, published by Professor Chan in January 2014.[8]

PISA's Response

Much of the Shanghai controversy involves empirical questions. Either hukou exists or it does not, either it has an effect on the migrant population in Shanghai or it does not. Surprisingly, the initial response of Andreas Schleicher, head of PISA, was to deny that hukou has any contemporary relevance in China, stating that "like many things in China, that has long changed." Those suggesting otherwise, Schleicher charged, were embracing old stereotypes.[9] That's a difficult stand to maintain given the enormous amount of press coverage and academic research recently published on hukou. A 2013 report from economists at OECD, the home of PISA, condemned the hukou system for barring migrants from high school. The report states,

> *To attend high school it is necessary to take an entrance examination and this must be taken in the locality of registration rather than the locality of residence. In Shanghai, migrant children can only attend vocational high schools. The Shanghai Education*

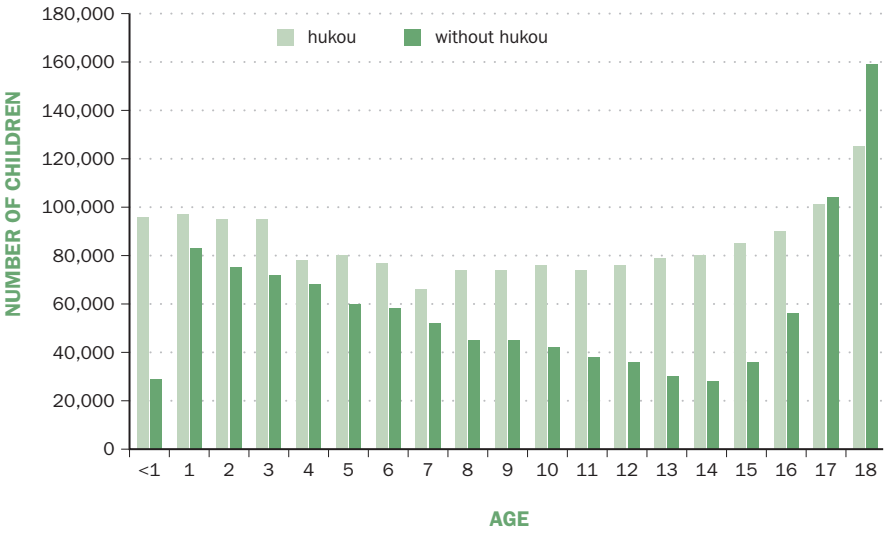
Committee justifies local high schools' refusal to admit the children of migrant workers on the grounds that "if we open the door to them, it would be difficult to shut in the future; local education resources should not be freely allocated to immigrant children." As a result, few migrant children attend general high schools and those who do return to their registration locality find it hard to adapt and often fail to complete the course.[10]

PISA's flat out denial of hukou did not last. It then shifted into an admission that, sure, hukou still exists, but its effects are not significant. Dr. Zhang Minxuan, President of Shanghai Normal University and the Shanghai coordinator for PISA, argued that because almost 30,000 migrants were part of the population sampled by PISA (which is in line with Professor Chan's number in Figure 1-1), no discrimination was occurring.[11] Marc Tucker and Andreas Schleicher published a response suggesting that questions about Shanghai were based on American chauvinism and envy of Shanghai's success.[12] All three cited China's one child policy as cause for the shortage of 15-year-olds, although the World Bank reports that the proportion of China's population from birth to 14-years-of-age (18%) is not extraordinary—exceeding countries such as Canada (16%), Germany (13%), Greece (15%), Japan (13%), and Korea (15%).[13] As is clear from Figure 1-1, the one child policy is not the problem. Hukou is the problem.

Mr. Schleicher and Mr. Tucker asked, who are you to throw stones at glass houses? Schleicher pointed out that the U.S. does not include Puerto Rico in its national PISA sample.[14] Marc Tucker drew a parallel between the Chinese hukou system and American treatment of racial minorities and socioeconomically disadvantaged students. As he put it,

Children in Shanghai, by Age & Hukou Status (2010 Census)

Fig. 1-1



Source: Kam Wing Chan, *Ming Pao,* January 2, 2014 (published in Chinese).

It is hard for me to distinguish the Hukou system from rural districts in the American south with white majorities who send their children to private academies and then put their fellow white citizens on the school board with instructions to impoverish the public schools serving African-American students in order to keep their taxes down, or northern communities which segregate poor and minority students into districts with no property wealth and high tax rates, while wealthy people are allowed to congregate in communities with low tax rates, gorgeous buildings and highly paid teachers.[15]

"If we open the door to [migrants], it would be difficult to shut in the future"

–The Shanghai Education Committee

PISA has held up Shanghai as a model of equity since 2009.

Schleicher, Tucker, and Zhang are honorable people. How could they defend the seemingly indefensible? Why would they be interested in downplaying the discriminatory effects of the hukou system, a policy condemned by groups such as Human Rights Watch and Amnesty International, a policy relic from China's authoritarian past that adversely affects hundreds of millions of migrant workers who have moved from the countryside to cities, a policy that separates parents from their children, and a policy that the Chinese government itself has promised to reform?

The answer is because PISA has held up Shanghai as a model of equity since 2009. PISA publications have never even mentioned hukou. It is embarrassing that an entire chapter is devoted to Shanghai in the OECD's *Strong Performers and Successful Reformers in Education* without a discussion of hukou. The following is said about the treatment of migrants: "In a way, Shanghai has established the notion that migrant children are 'our children' and works constructively to include them in its educational development."[16] Contrast such magnanimous sentiments with the quotation above that defends shutting out migrants.

Recommendations for the Governance of PISA

The integrity of PISA is at risk. The following three reforms would help restore PISA's legitimacy. The PISA Governing Board (PGB) should immediately take action to implement them.

1. Ensure the Independence of the PISA Governing Board (PGB) Policy relevance has become PISA's overarching objective, which has led to excessive and unfounded policy recommendations. It has also created an inherent conflict of interest. The OECD is a quasi-governmental body; thus, PISA is a creature of government. The PGB is made up of representatives from the governments that take part in PISA. They are not disinterested in the results.

Shanghai is a prime example of the conflict. Here is Marc Tucker describing Dr. Zhang, who, as pointed out above, coordinates the administration of PISA in Shanghai: "From 2004 through 2011, Zhang was the Vice-Director General of the Shanghai Municipal Education Committee, and, in that capacity, in charge of planning many of the education reforms for which Shanghai has since become famous."[17]

Reforms in Shanghai became famous because of PISA. Shanghai's PISA test is run by the same official who advanced the reforms. That sure sounds like a conflict of interest. Moreover, the Shanghai Municipal Education Committee is responsible for enforcing the hukou system's education restrictions. Dr. Zhang served in a leadership capacity on that body. No wonder PISA documents are silent on the negative effects of hukou. The larger lesson concerning governance and policy recommendations applies to every PISA participant. Governments demand policy guidance from an assessment that looks at how well the policies that they themselves have enacted are functioning. It is difficult to be an impartial referee while also playing in the game.

Checks and balances need to be built into the PISA governance structure. It is unreasonable to expect national governments, who pay for international assessments, not to play a role in their governance. But that authority should be shared with educators, statisticians, policy analysts, and other independent experts from education's disciplinary fields. The consumers of test data extend beyond government. In the U.S., for example, the National Assessment

of Educational Progress (NAEP) is governed by a board that includes citizens from many walks of life and whose independence from the federal government is structurally guarded. In 1999, when Vice President Al Gore, who happened to be a prospective candidate for President at the time, appeared at a press conference to release NAEP data, the inappropriateness of the event stirred howls of outrage.[18]

The Trends in International Mathematics and Science Study (TIMSS) tests are governed by the General Assembly of the International Association for the Evaluation of Educational Achievement (IEA), which includes non-governmental representatives. The author of this report served as the U.S. representative to the IEA General Assembly from 2004–2012. Before him, the U.S. representatives to the IEA were from the research community (e.g., Richard M. Wolf) or worked in organizations representing state and local authorities (e.g., Gordon M. Ambach of the Council of Chief State School Officers). The PGB needs a similarly diverse membership.

2. Separate the Policy Recommenders from the Data Collectors In the U.S., the National Assessment of Educational Progress administers a national test and reports results. Analysis is primarily conducted by independent consumers of the data. Officials from NAEP or the National Center for Education Statistics (NCES) do not speculate as to why particular states score at the top of NAEP league tables or recommend that states adopt particular policies to improve their performance. It is strange that the U.S. participates in an international test that violates the constraints it imposes on its own national assessment.

The same subunit of OECD plans and administers PISA, analyzes the data, and makes policy recommendations. PISA data releases are accompanied by thematic volumes. The 2012 data, for example, were joined by volumes on equity, student engagement, and the characteristics of successful schools. The title of the 2012 volume on school characteristics reveals its ambitions: "What Makes Schools Successful?"[19]

These volumes are jam-packed with policy recommendations. Almost all of the recommendations are based on analyses of cross-sectional data. Skilled policy analysts are cautious in making policy recommendations based on cross-sectional data because they provide weak evidence for policy guidance.

A good example of potential pitfalls can be seen in PISA's interpretation of the data on pre-primary education. The PISA score difference between students who, at age 15, report that they had attended pre-primary school and those who did not is large and statistically significant, 53 points before controlling for SES and 31 points after SES controls are employed. The standard deviation of PISA being 100, the gaps equate to about one-half and one-third of a standard deviation. PISA recommends more and earlier pre-primary education, part of the PISA education agenda promoted since the first PISA assessment in 2000.[20]

Asking 15-year-olds whether they attended pre-primary school is not strong evidence on which to recommend more extensive pre-primary education. Controlling for SES does not control for all of the unmeasured variables that may be driving higher math scores at age 15. Parents who care more about their toddler's academic success in the future, for example, may be more inclined to enroll their children in pre-primary schools—and may be more likely to emphasize mathematics when their children are teenagers. Moreover, the PISA data collectors now have taken a political position that they may feel

PISA officials should not advocate education policies because the data that they collect reflect on the wisdom of those policies.

Who in China took the PISA in 2009 and 2012 remains a mystery.

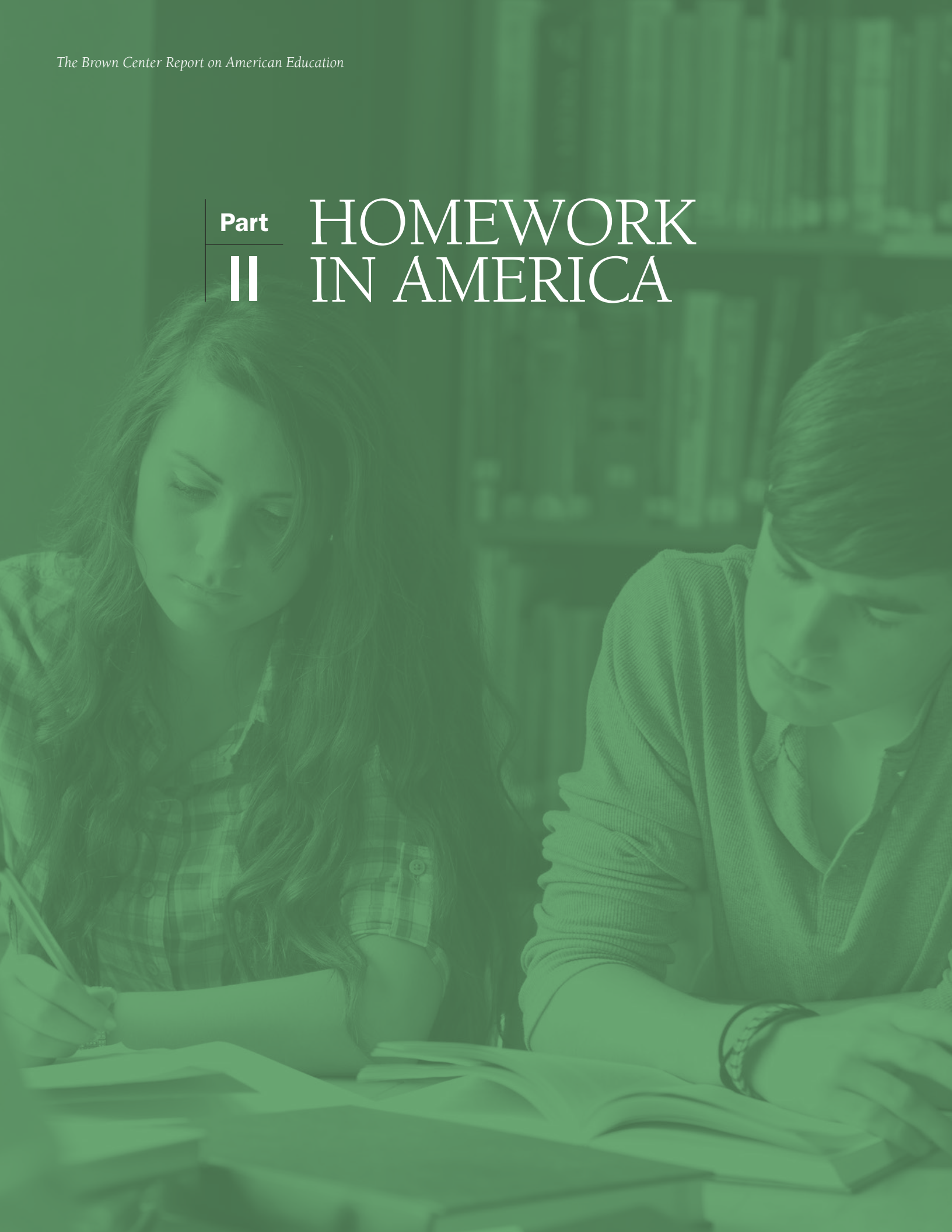
compelled to defend. PISA officials should not advocate education policies because the data that they collect reflect on the wisdom of those policies.

3. Full Transparency on the Expansion of PISA in China PISA officials have announced plans to expand PISA testing to other provinces in China besides Shanghai in 2015. The expansion affords PISA an opportunity to dispel the secrecy surrounding the conduct of the assessment in China. Since Shanghai first participated in 2009, Andreas Schleicher has made several references in the press to PISA data collected in other Chinese provinces.[21] But the data have never been released.

Who conducted the tests? Mr. Schleicher originally made statements such as, "We have actually done PISA in 12 of the provinces in China," as he told the *Financial Times* on December 7, 2010.[22] After being challenged to release the data, Mr. Schleicher changed the story in an interview on *Daybreak Asia* to "the Chinese authorities themselves have been experimenting with PISA in other provinces, but, you know, those data are not representative so we don't use them in our comparisons."[23] Who in China took the PISA in 2009 and 2012 remains a mystery. Whether the Chinese government or OECD decided which test scores could be released also remains a mystery. Referring to 2009 PISA data from rural Chinese provinces, a 2012 *BBC* report stated that "The Chinese government has so far not allowed the OECD to publish the actual data."[24]

All of this needs to be cleared up before the 2015 PISA. The PGB should appoint an independent panel to investigate the conduct of PISA in China, collect relevant evidence on agreements OECD made with the Chinese government as to its participation, and release all data collected in 2009 and 2012 to the public domain. Otherwise, future PISA activities in China operate under a cloud of suspicion and doubt.

Part II HOMEWORK IN AMERICA



HOMEWORK! THE TOPIC, NO, JUST THE WORD ITSELF, SPARKS controversy. It has for a long time. In 1900, Edward Bok, editor of the *Ladies Home Journal*, published an impassioned article, "A National Crime at the Feet of Parents," accusing homework of destroying American youth. Drawing on the theories of his fellow educational progressive, psychologist G. Stanley Hall (who has since been largely discredited), Bok argued that study at home interfered with children's natural inclination towards play and free movement, threatened children's physical and mental health, and usurped the right of parents to decide activities in the home.

The *Journal* was an influential magazine, especially with parents. An anti-homework campaign burst forth that grew into a national crusade.[25] School districts across the land passed restrictions on homework, culminating in a 1901 statewide prohibition of homework in California for any student under the age of 15. The crusade would remain powerful through 1913, before a world war and other concerns bumped it from the spotlight. Nevertheless, anti-homework sentiment would remain a touchstone of progressive education throughout the twentieth century. As a political force, it would lie dormant for years before bubbling up to mobilize proponents of free play and "the whole child." Advocates would, if educators did not comply, seek to impose homework restrictions through policy making.

Our own century dawned during a surge of anti-homework sentiment. From 1998 to 2003, *Newsweek*, *TIME*, and *People*, all major national publications at the time, ran cover stories on the evils of homework. *TIME*'s 1999 story had the most provocative title, "The Homework Ate My Family: Kids Are Dazed, Parents Are Stressed, Why Piling On Is Hurting Students." *People*'s 2003 article offered a call to arms: "Overbooked: Four Hours of Homework for a Third Grader? Exhausted Kids (and Parents) Fight Back." Feature stories about students laboring under an onerous homework burden ran in newspapers from coast to coast. Photos of angst

ridden children became a journalistic staple.

The 2003 Brown Center Report on American Education included a study investigating the homework controversy. Examining the most reliable empirical evidence at the time, the study concluded that the dramatic claims about homework were unfounded. An overwhelming majority of students, at least two-thirds, depending on age, had an hour or less of homework each night. Surprisingly, even the homework burden of college-bound high school seniors was discovered to be rather light, less than an hour per night or six hours per week. Public opinion polls also contradicted the prevailing story. Parents were not up in arms about homework. Most said their children's homework load was about right. Parents wanting more homework out-numbered those who wanted less.

Now homework is in the news again. Several popular anti-homework books fill store shelves (whether virtual or brick and mortar).[26] The documentary *Race to Nowhere* depicts homework as one aspect of an overwrought, pressure-cooker school system that constantly pushes students to perform and destroys their love of learning. The film's website claims over 6,000 screenings in more than 30 countries. In 2011, the *New York Times* ran a front page article about the homework restrictions adopted by schools in Galloway, NJ, describing "a wave of districts across the nation trying to remake homework amid concerns that high stakes testing and competition for college have fueled a nightly grind that is stressing out children and depriving them of play and rest, yet doing little to raise achievement, especially in elementary grades." In the article, Vicki Abeles, the director of *Race to Nowhere*, invokes the indictment of homework lodged a century ago, declaring, "The presence of homework is negatively affecting the health of our young people and the quality of family time."[27]

A petition for the National PTA to adopt "healthy homework guidelines" on change.org currently has 19,000 signatures. In September 2013, *Atlantic* featured an article, "My Daughter's Homework is Killing Me," by a Manhattan writer who joined his middle school daughter in doing her homework for a week. Most nights the homework took more than three hours to complete.

The Current Study

A decade has passed since the last Brown Center Report study of homework, and it's time for an update. How much homework do American students have today? Has the homework burden increased, gone down, or remained about the same? What do parents think about the homework load?

A word on why such a study is important. It's not because the popular press is creating a fiction. The press accounts are built on the testimony of real students and real parents, people who are very unhappy with the amount of homework coming home from school. These unhappy people are real—but they also may be atypical. Their experiences, as dramatic as they are, may not represent the common experience of American households with school-age children. In the analysis below, data are analyzed from surveys that are methodologically designed to produce reliable information about the experiences of all Americans. Some of the surveys have existed long enough to illustrate meaningful trends. The question is whether strong empirical evidence confirms the anecdotes about overworked kids and outraged parents.

NAEP Data

Data from the National Assessment of Educational Progress (NAEP) provide a good

Homework is in the news again. Several popular anti-homework books fill store shelves.

The proportion of [13-year-old] students with the heaviest load, more than two hours of homework, slipped from 9% in 1984 to 7% in 2012.

look at trends in homework for nearly the past three decades. Table 2-1 displays NAEP data from 1984–2012. The data are from the long-term trend NAEP assessment's student questionnaire, a survey of homework practices featuring both consistently-worded questions and stable response categories. The question asks: "How much time did you spend on homework yesterday?" Responses are shown for NAEP's three age groups: 9, 13, and 17.[28]

Today's youngest students seem to have more homework than in the past. The first three rows of data for age 9 reveal a shift away from students having no homework, declining from 35% in 1984 to 22% in 2012. A slight uptick occurred from the low of 18% in 2008, however, so the trend may be abating. The decline of the "no homework" group is matched by growth in the percentage of students with less than an hour's worth, from 41% in 1984 to 57% in 2012. The share of students with one to two hours of homework changed very little over the entire 28 years, comprising 12% of students in 2012. The group with the heaviest load, more than two hours of homework, registered at 5% in 2012. It was 6% in 1984.

The amount of homework for 13-year-olds appears to have lightened slightly. Students with one to two hours of homework declined from 29% to 23%. The next category down (in terms of homework load), students with less than an hour, increased from 36% to 44%. One can see, by combining the bottom two rows, that students with an hour or more of homework declined steadily from 1984 to 2008 (falling from 38% to 27%) and then ticked up to 30% in 2012. The proportion of students with the heaviest load, more than two hours, slipped from 9% in 1984 to 7% in 2012 and ranged between 7–10% for the entire period.

For 17-year-olds, the homework burden has not varied much. The percentage of students with no homework has increased from 22% to 27%. Most of that gain occurred in the 1990s. Also note that the percentage of 17-year-olds who had homework but did not do it was 11% in 2012, the highest for the three NAEP age groups. Adding that number in with the students who didn't have homework in the first place means that more than one-third of 17-year-olds (38%) did no homework on the night in question in 2012. That compares with 33% in 1984. The segment of the 17-year-old population with more than two

Students Were Asked
How much time did you spend on homework yesterday?
(percent of students)

Table 2-1

Age 9

	1984	1992	1999	2008	2012
None Assigned	35	32	26	18	22
Did Not Do It	4	4	4	5	4
< 1 hr.	41	47	53	60	57
1-2 hrs.	13	12	12	12	12
>2hrs.	6	5	5	5	5

Age 13

	1984	1992	1999	2008	2012
None Assigned	22	21	24	23	21
Did Not Do It	4	4	5	7	5
< 1 hr.	36	36	37	43	44
1-2 hrs.	29	29	26	21	23
>2hrs.	9	10	8	6	7

Age 17

	1984	1992	1999	2008	2012
None Assigned	22	22	26	28	27
Did Not Do It	11	12	13	12	11
< 1 hr.	26	29	26	27	26
1-2 hrs.	27	25	23	22	23
>2hrs.	13	11	12	10	13

Source: NAEP Data Explorer, long term trend reading data for ages 9, 13, and 17 (Item B001701).

hours of homework, from which legitimate complaints of being overworked might arise, has been stuck in the 10%–13% range.

The NAEP data point to four main conclusions:

1. With one exception, the homework load has remained remarkably stable since 1984.
2. The exception is 9-year-olds. They have experienced an increase in homework, primarily because many students who once did not have any now have some. The percentage of 9-year-olds with no homework fell by 13 percentage points, and the percentage with less than an hour grew by 16 percentage points.
3. Of the three age groups, 17-year-olds have the most bifurcated distribution of the homework burden. They have the largest percentage of kids with no homework (especially when the homework shirkers are added in) and the largest percentage with more than two hours.
4. NAEP data do not support the idea that a large and growing number of students have an onerous amount of homework. For all three age groups, only a small percentage of students report more than two hours of homework. For 1984–2012, the size of the two hours or more groups ranged from 5–6% for age 9, 6–10% for age 13, and 10–13% for age 17.

Note that the item asks students how much time they spent on homework "yesterday." That phrasing has the benefit of immediacy, asking for an estimate of precise, recent behavior rather than an estimate of general behavior for an extended, unspecified period. But misleading responses could be generated if teachers lighten the homework of NAEP participants on the night before the NAEP test is given. That's possible.[29] Such skewing would not affect trends if it stayed about the same over time and in the same direction (teachers assigning less homework than usual on the day before NAEP). Put another way, it would affect estimates of the amount of homework at any single point in time but not changes in the amount of homework between two points in time.

A check for possible skewing is to compare the responses above with those to another homework question on the NAEP questionnaire from 1986–2004 but no longer in use.[30] It asked students, "How much time do you usually spend on homework each day?" Most of the response categories have different boundaries from the "last night" question, making the data incomparable. But the categories asking about no homework are comparable. Responses indicating no homework on the "usual" question in 2004 were: 2% for 9-year-olds, 5% for 13-year-olds, and 12% for 17-year-olds. These figures are much less than the ones reported in Table 2-1 above. The "yesterday" data appear to overstate the proportion of students typically receiving no homework.

The story is different for the "heavy homework load" response categories. The "usual" question reported similar percentages as the "yesterday" question. The categories representing the most amount of homework were "more than one hour" for age 9 and "more than two hours" for ages 13 and 17. In 2004, 12% of 9-year-olds said they had more than one hour of daily homework, while 8% of 13-year-olds and 12% of 17-year-olds said they had more than two hours. For all three age groups, those figures declined from 1986 to 2004. The decline for age 17 was quite large, falling from 17% in 1986 to 12% in 2004.

The bottom line: regardless of how the question is posed, NAEP data do not support the view that the homework burden is growing, nor do they support the belief that the proportion of students with a lot of homework has increased in recent years.

NAEP data do not support the view that the homework burden is growing … or the belief that the proportion of students with a lot of homework has increased.

Only 38.4% of [college freshmen] said they spent at least six hours per week studying or doing homework as seniors in high school."

The proportion of students with no homework is probably under-reported on the long-term trend NAEP. But the upper bound of students with more than two hours of daily homework appears to be about 15%—and that is for students in their final years of high school.

College Freshmen Look Back

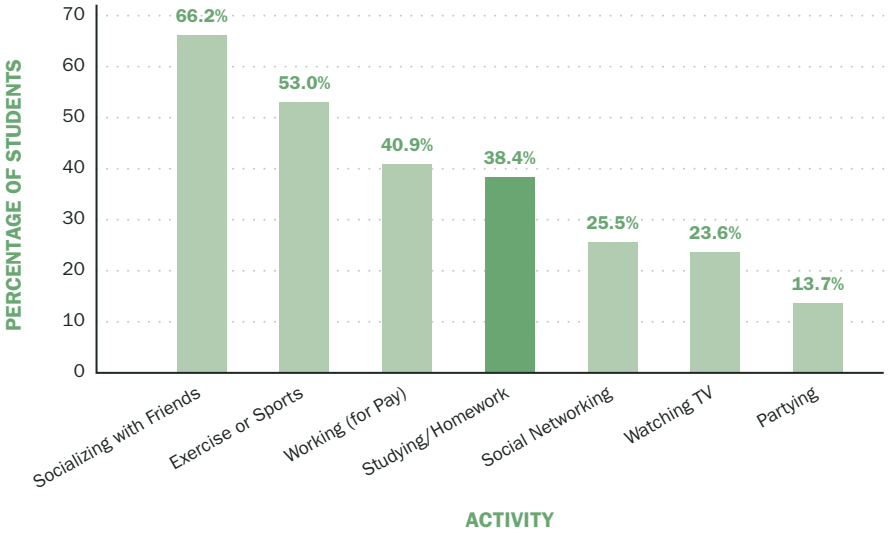
There is another good source of information on high school students' homework over several decades. The Higher Education Research Institute at UCLA conducts an annual survey of college freshmen that began in 1966. In 1986, the survey started asking a series of questions regarding how students spent time in the final year of high school. Figure 2-1 shows the 2012 percentages for the dominant activities. More than half of college freshmen say they spent at least six hours per week socializing with friends (66.2%) and exercising/sports (53.0%). About 40% devoted that much weekly time to paid employment.

Homework comes in fourth pace. Only 38.4% of students said they spent at least six hours per week studying or doing homework. When these students were high school seniors, it was not an activity central to their out of school lives. That is quite surprising. Think about it. The survey is confined to the nation's best students, those attending college. Gone are high school dropouts. Also not included are students who go into the military or attain full-time employment immediately after high school. And yet only a little more than one-third of the sampled students devoted more than six hours per week to homework and studying when they were on the verge of attending college.

Another notable finding from the UCLA survey is how the statistic is trending (see Figure 2-2). In 1986, 49.5% reported spending six or more hours per week studying and doing homework. By 2002, the proportion had dropped to 33.4%. In 2012, as noted in Figure 2-1, the statistic had bounced off the historical lows to reach 38.4%. It is slowly rising but still sits sharply below where it was in 1987.

Students Spending Six or More Hours Per Week on Activities in their Last Year of High School (percent of students)

Figure 2-1



Source: John H. Pryor, K. Eagan, L. Palucki Blake, S. Hurtado, J. Berdan, and M. Case. "The American Freshman: National Norms Fall 2012," Higher Education Research Institute, UCLA, 2012.

What Do Parents Think?

MetLife has published an annual survey of teachers since 1984. In 1987 and 2007, the survey included questions focusing on homework and expanded to sample both parents and students on the topic. Data are broken out for secondary and elementary parents and for students in grades 3–6 and grades 7–12 (the latter not being an exact match with secondary parents because of K–8 schools).

Table 2-2 shows estimates of homework from the 2007 survey. Respondents were asked to estimate the amount of homework on a typical school day (Monday–

Friday). The median estimate of each group of respondents is shaded. As displayed in the first column, the median estimate for parents of an elementary student is that their child devotes about 30 minutes to homework on the typical weekday. Slightly more than half (52%) estimate 30 minutes or less; 48% estimate 45 minutes or more. Students in grades 3–6 (third column) give a median estimate that is a bit higher than their parents' (45 minutes), with almost two-thirds (63%) saying 45 minutes or less is the typical weekday homework load.

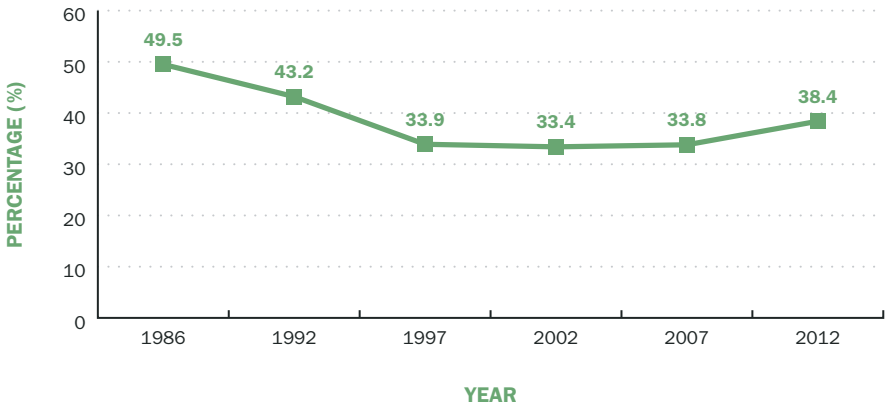
One hour of homework is the median estimate for both secondary parents and students in grade 7–12, with 55% of parents reporting an hour or less and about two-thirds (67%) of students reporting the same. As for the prevalence of the heaviest homework loads, 11% of secondary parents say their children spend more than two hours on weekday homework, and 12% is the corresponding figure for students in grades 7–12.

The MetLife surveys in 1987 and 2007 asked parents to evaluate the amount and quality of homework. Table 2-3 displays the results. There was little change over the two decades separating the two surveys. More than 60% of parents rate the amount of homework as good or excellent, and about two-thirds give such high ratings to the quality of the homework their children are receiving. The proportion giving poor ratings to either the quantity or quality of homework did not exceed 10% on either survey.

Parental dissatisfaction with homework comes in two forms: those who feel schools give too much homework and those who feel schools do not give enough. The current wave of journalism about unhappy parents is dominated by those who feel schools give too much homework. How big is this group? Not very big (see Figure 2-3). On the MetLife survey, 60% of parents

High School Seniors Spending Six or More Hours Per Week Studying or Doing Homework, 1986–2012 (percent of students)

Figure 2-2



Source: Compiled by author from The American Freshman: National Norms from 1986, 1992, 1997, 2002, 2007, and 2012 (Higher Education Research Institute, UCLA).

Parent and Student Estimates of Homework on Typical School Day (percent of respondents)

Table 2-2

	Parents		Students	
	Elementary	Secondary	Grades 3–6	Grades 7–12
None	2	4	2	8
5 mins	5	*	5	4
15 mins	16	3	15	10
30 mins	29	11	23	17
45 mins	12	12	18	10
1 hour	17	25	18	18
1.5 hours	10	14	9	11
2 hours	5	15	5	9
2.5 hours	1	6	2	4
3 hours or more	3	5	2	8

Note: Shaded Cells indicate location of sample median. Responses of "Not Sure" not reported.

* N/A

Source: Figures 2.4 and 2.8, *Met Life Survey of the American Teacher, The Homework Experience*, MetLife, 2007.

Polls show that parents who want less homework are outnumbered by parents who want more homework.

felt schools were giving the right amount of homework, 25% wanted more homework, and only 15% wanted less.

National surveys on homework are infrequent, but the 2006–2007 period had more than one. A poll conducted by Public Agenda in 2006 reported similar numbers as the MetLife survey: 68% of parents describing the homework load as "about right," 20% saying there is "too little homework," and 11% saying there is "too much homework." A 2006 AP–AOL poll found the highest percentage of parents reporting too much homework, 19%. But even in that poll, they were outnumbered by parents believing there is too little homework (23%), and a clear majority (57%) described the load as "about right." A 2010 local survey of Chicago parents conducted by the *Chicago Tribune* reported figures similar to those reported above: approximately two-thirds of parents saying their children's homework load is "about right," 21% saying it's not enough, and 12% responding that the homework load is too much.

Parent Ratings of Homework Amount and Quality, 2007 & 1987 (percent of respondents)

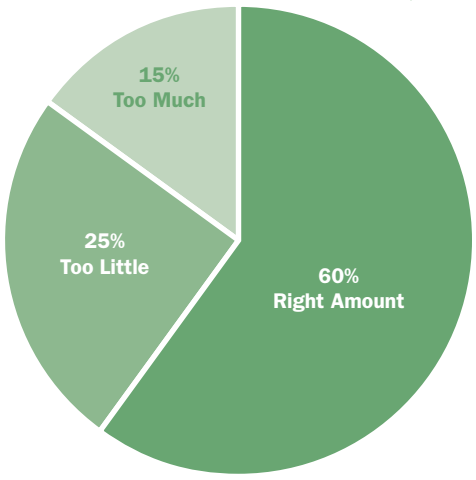
Table 2-3

		Excellent	Good	Fair	Poor	Not Sure
Amount	2007	15	46	30	9	–
	1987	16	47	24	9	4
Quality	2007	19	48	28	5	–
	1987	17	49	22	7	4

Sources: *Met Life Survey of the American Teacher, The Homework Experience,* MetLife, 1987
Met Life Survey of the American Teacher, The Homework Experience, MetLife, 2007.

Parents' View of Homework Load, 2007

Figure 2-3



Source: Figure 6.14, *Met Life Survey of the American Teacher, The Homework Experience*, MetLife, 2007.

Summary and Discussion

In recent years, the press has been filled with reports of kids over-burdened with homework and parents rebelling against their children's oppressive workload. The data assembled above call into question whether that portrait is accurate for the typical American family. Homework typically takes an hour per night. The homework burden of students rarely exceeds two hours a night. The upper limit of students with two or more hours per night is about 15% nationally—and that is for juniors or seniors in high school. For younger children, the upper boundary is about 10% who have such a heavy load. Polls show that parents who want less homework range from 10%–20%, and that they are outnumbered—in every national poll on the homework question—by parents who want more homework, not less. The majority of parents describe their children's homework burden as about right.

So what's going on? Where are the homework horror stories coming from?

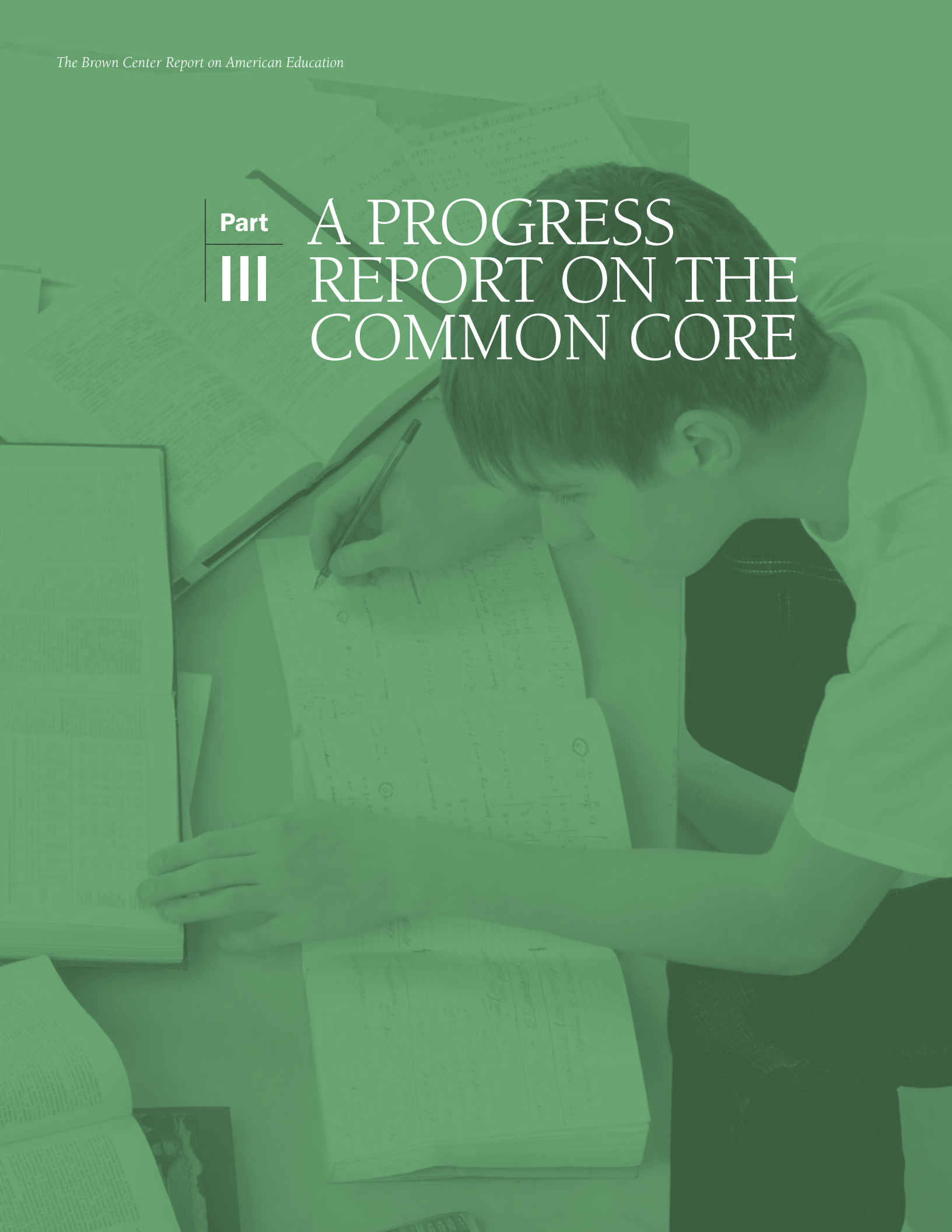
The MetLife survey of parents is able to give a few hints, mainly because of several questions that extend beyond homework to other aspects of schooling. The belief that homework is burdensome is more likely

held by parents with a larger set of complaints and concerns. They are alienated from their child's school. About two in five parents (19%) don't believe homework is important. Compared to other parents, these parents are more likely to say too much homework is assigned (39% vs. 9%), that what is assigned is just busywork (57% vs. 36%), and that homework gets in the way of their family spending time together (51% vs. 15%). They are less likely to rate the quality of homework as excellent (3% vs. 23%) or to rate the availability and responsiveness of teachers as excellent (18% vs. 38%).[31]

They can also convince themselves that their numbers are larger than they really are. Karl Taro Greenfeld, the author of the *Atlantic* article mentioned above, seems to fit that description. "Every parent I know in New York City comments on how much homework their children have," Mr. Greenfeld writes. As for those parents who do not share this view? "There is always a clique of parents who are happy with the amount of homework. In fact, they would prefer *more*. I tend not to get along with that type of parent."[32]

Mr. Greenfeld's daughter attends a selective exam school in Manhattan, known for its rigorous expectations and, yes, heavy homework load. He had also complained about homework in his daughter's previous school in Brentwood, CA. That school was a charter school. After Mr. Greenfeld emailed several parents expressing his complaints about homework in that school, the school's vice-principal accused Mr. Greenfeld of cyberbullying. The lesson here is that even schools of choice are not immune from complaints about homework.

The homework horror stories need to be read in a proper perspective. They seem to originate from the very personal discontents of a small group of parents. They do not reflect the experience of the average family with a school-age child. That does not diminish these stories' power to command the attention of school officials or even the public at large. But it also suggests a limited role for policy making in settling such disputes. Policy is a blunt instrument. Educators, parents, and kids are in the best position to resolve complaints about homework on a case by case basis. Complaints about homework have existed for more than a century, and they show no signs of going away.



Part III A PROGRESS REPORT ON THE COMMON CORE

WILLIAM H. SCHMIDT OF MICHIGAN STATE UNIVERSITY presented research on the Common Core State Standards (CCSS) for Mathematics at the National Press Club on May 3, 2012.[33] A paper based on the same research, co-authored with Richard T. Houang, was published in *Educational Researcher* in October 2012.[34] Schmidt and Houang's study (also referred to as the "MSU study" below) was important for endorsing CCSS's prospective effectiveness at a time when debate on the CCSS was beginning to heat up. Opponents of the Common Core had criticized the CCSS for lacking empirical support.

The MSU study showed that states with math standards similar to the Common Core, after controlling for other potential influences, registered higher NAEP scores in 2009 than states with standards divergent from the CCSS. The implication was that the math standards of CCSS would boost state math performance on NAEP.

Is there reason to believe that projection will become reality? In this section of the Brown Center Report, a two-part investigation attempts to answer that question. First, the ratings of state standards provided by Schmidt and Houang's study are examined using NAEP data that have been collected since their study was completed. The central question is whether the MSU ratings predict progress on NAEP from 2009–2013. Second, a new analysis is presented, independent from the MSU ratings, comparing the NAEP gains of states with varying degrees of CCSS implementation. The two analyses offer exploratory readings of how the Common Core is affecting achievement so far.

Background

Schmidt and Houang used state NAEP scores on the 2009 eighth grade math assessment to model the potential effectiveness of the CCSS. They first developed a scale to rate the degree of congruence of each state's standards with the CCSS. The ratings were based on earlier work also conducted by Schmidt and his colleagues at MSU. That work made a lasting and important contribution to curriculum studies by attempting to represent the quality of

curriculum standards—both international and domestic—in a quantitative form.[35] The key dimensions measured in the MSU ratings are focus and coherence. Focus refers to limiting topics in the math curriculum to the most important topics and teaching them in depth. Coherence refers to organizing topics in a manner that reflects the underlying structure of mathematics, allowing knowledge and skills to build sequentially.

In the National Press Club talk, Schmidt presented a chart showing how the states fell on the congruence measure (see Table 3-1). Alabama, Michigan, California, and the others at the top of the scale had standards most like the CCSS math standards. Arizona, Nevada, Iowa and those at the bottom of the scale had standards that diverged from the CCSS.

Table 3-1 includes a categorical variable (1–5) for the five congruency ratings. The MSU authors used the continuous form of the congruence ratings along with demographic covariates in a regression equation that predicted state NAEP scores. The congruence rating was statistically insignificant. No relationship to achievement was uncovered. An analysis of residuals, however, revealed two distinct sets of states (referred to as "Group A" and "Group B"). (Key differences between the two groups are discussed below.) Regression equations incorporating membership in these two groups did produce statistically significant coefficients for the congruence rating.

Figure 3-1, reproduced from the *Educational Researcher* article, clearly shows two upward sloping regression lines. The MSU authors concluded that it was time to end the debate over the wisdom of the Common Core and that the CCSS in math "deserve to be seriously implemented."[36]

Existing State Standards' Consistency with the Common Core State Standards for Mathematics

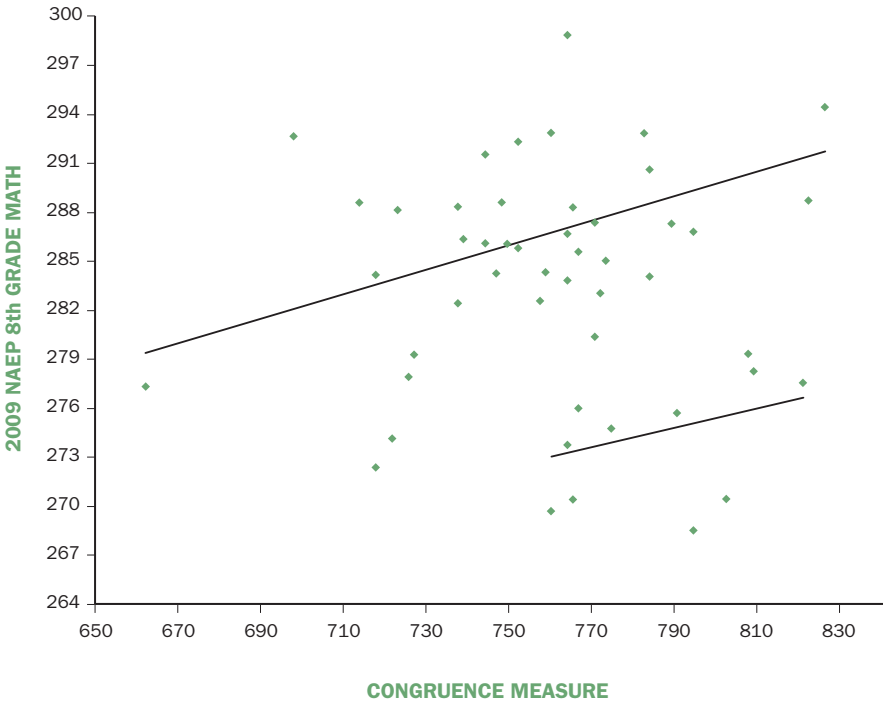
Table 3-1

Most Like CCSS	5	Alabama, California, Florida, Georgia, Indiana, Michigan, Minnesota, Mississippi, Oklahoma, Washington
	4	Idaho, North Dakota, Oregon, South Dakota, Tennessee, Utah
	3	Alaska, Arkansas, Colorado, Delaware, Hawaii, Massachusetts, New Mexico, New York, North Carolina, Ohio, Pennsylvania, South Carolina, Texas, Vermont, West Virginia
	2	Connecticut, Illinois, Maine, Maryland, Missouri, Montana, Nebraska, New Hampshire, Virginia, Wyoming
Least Like CCSS	1	Arizona, Iowa, Kansas, Kentucky, Louisiana, Nevada, New Jersey, Rhode Island, Wisconsin

Source: William H. Schmidt (2012, May). *The Common Core State Standards for Mathematics*. Presented at the National Press Club, Washington, D.C. Categorical labels 1–5 assigned by author.

Scatter Plot with Estimated Regression Lines for Groups A and B Relating Congruence to 2009 National Assessment of Educational Progress (NAEP)

Fig. 3-1



Note: Mississippi was excluded from the scatter plot.

Source: William H. Schmidt and Richard T. Houang. "Curricular Coherence and the Common Core State Standards for Mathematics," *Educational Researcher* November 2012 41: 294–308, copyright © 2012, SAGE Publications. Reprinted by permission of SAGE Publications.

States with standards most different from the CCSS gained the most on NAEP.

Examining NAEP Gains with the MSU Ratings

NAEP scores for 2011 and 2013 have been released since the Schmidt and Houang study. These scores offer the opportunity to update Schmidt and Houang's findings. They also allow a check of the study's most important policy lesson, that states adopting the CCSS in math could expect an increase in their eighth grade NAEP math scores. Examining gain scores—specifically, the change in state scores since 2009—provides a way to evaluate the predictive capacity, at least in the short run, of the 2009 MSU analysis. By relying on cross-sectional data, NAEP scores from a single point in time, the Schmidt and Houang analysis helps to explain the performance of states in 2009. But states adopt policies with an eye toward future results. Did the states with standards most like the CCSS in 2009 continue to make the greatest gains in later years? Gain score analysis also possesses a technical advantage. It is generally superior to cross-sectional analysis in controlling for omitted variables that may influence achievement by "baking them into the cake" at baseline.[37]

State NAEP Changes, by MSU's Rating of Congruence with CCSS (in scale score points, 2009–2013)

Table 3-2

MSU Rating	2009–2011	2011–2013	2009–2013
5 (n=10)	1.00	0.94	1.94
4 (n=6)	–0.91	0.10	–0.81
3 (n=15)	1.59	0.21	1.80
2 (n=10)	0.23	0.06	0.29
1 (n=9)	1.91	0.34	2.25
All (n=50)	0.96	0.34	1.30

Source: Generated from NAEP Data Explorer data. MSU Rating obtained from William H. Schmidt's May 2012 presentation to the National Press Club, Washington, D.C. Categorical labels 1–5 assigned by author.

Tables 3-2, 3-3, and 3-4 report the average gains of states expressed as changes in scale score points on the eighth grade NAEP math assessment. The states are grouped by their MSU rating. Bear in mind that the 2009 MSU ratings were assigned based on the math standards then in place. States with a "5" had math standards most similar to the CCSS. States with a "1" had math standards most divergent from the CCSS.

Table 3-2 reveals no systematic relationship between the states' MSU ratings and changes in NAEP from 2009–2013. Indeed, states with standards most different from the CCSS (rated 1) gained the most on NAEP (2.25). States with standards most like the CCSS scored the next largest gains (1.94); and states with a 4 rating (second most similar group to the CCSS) lost ground, declining -0.81. The data are lumpy, so whether a positive relationship is expected (i.e., states scoring 5 should make the greatest gains, 4 the next greatest gains, and so forth) or a negative relationship (states scoring 1 should make the greatest gains because they have the most to gain from adopting CCSS, states scoring 2 have the next most to gain, etc.), no statistical relationship is evident. No linear pattern emerges across the categories.

What about the two time intervals, 2009–2011 and 2011–2013? NAEP scores are more stable over longer periods of time so the four year interval is probably a preferable indicator. In addition, a clear point of demarcation does not exist for when an old set of standards ends and a new set begins. Nevertheless, let's consider how the CCSS unfolded to guide the consideration of the data by different time periods.

The 2009–2011 interval should probably receive the closest scrutiny in probing for a correlation of state achievement with 2009 standards. Those standards were still operational from 2009–2011. The states rated

"1" notched particularly strong gains (1.91) during this period. States rated "4" actually declined (-0.91). That is not what one would expect if the MSU ratings accurately reflected the quality of 2009 standards.

The 2011–2013 interval should represent the strongest early indicator of gains after adopting the CCSS. Forty-five states had adopted the CCSS math standards by 2011. In a survey of state officials in 2011, most declared that they had begun the implementation process (progress in implementation receives explicit attention below).[38] The gains for this interval might be expected to be inversely related to the MSU ratings, with larger gains coming from the states rated "1." They were making the most dramatic curricular changes and should experience the most growth that accrues from adopting the CCSS. That expectation isn't met either. States with a "5" made the largest gains (0.94); however, the second largest gains were recorded by the states with a "1" rating (0.34).

Recall that Schmidt and Houang did not find a significant relationship until they divided the states into two groups, Group A and Group B. Group A consists of 37 states and Group B has 13 states. The groups are quite different demographically. More than half of the Group B states are Southern. They have lower per capita wealth and serve a greater proportion of black and Hispanic students. They receive more federal funding than Group A states. They also scored about 14.67 scale score points lower than Group A states on the 2009 NAEP. Schmidt and Houang speculate that the states in Group B, despite many having high quality math standards, faced a more difficult implementation environment because of demographic challenges and resource constraints.

Tables 3-3 and 3-4 disaggregate the gains by these two groups. Table 3-3 examines the A group. The NAEP changes generally contradict the MSU ratings. From 2009–2013, the states with the weakest congruence with CCSS made the greatest gains (2.25). The changes from 2009–2011 are the most glaring. States with the strongest ratings (the 5's) lost ground (-0.64), and the states rated "1" scored gains (1.91). Note, though, that some of the ratings groups have become quite small (only three states in Group A have a "5" rating), so these figures must be taken with several grains of salt. Also observe that all of the states with ratings of "1" or "2" belong to group A. Consequently, the results for these states in Table 3-3 are the same as in Table 3-2.

Changes in NAEP Scores (in scale score points), Group A States (37 states)

Table 3-3

MSU Rating	2009–2011	2011–2013	2009–2013
5 (n=3)	−0.64	1.42	0.78
4 (n=5)	−0.95	−0.61	−1.56
3 (n=10)	0.85	0.25	1.10
2 (n=10)	0.23	0.06	0.29
1 (n=9)	1.91	0.34	2.25
All (n=37)	0.58	0.20	0.77

Source: Generated from NAEP Data Explorer data. MSU Rating obtained from William H. Schmidt's May 2012 presentation to the National Press Club, Washington, D.C. Categorical labels 1–5 assigned by author.

Table 3-4 examines the states in Group B. Note that the ratings divide what is already a small group, 13 states, into even smaller groups. The states rated "5" registered the smallest gain (2.44) of the ratings groups for 2009–2013. As a whole, from 2009–2013 the Group B states made larger gains than the Group A states (2.78 vs. the 0.77 reported in Table 3-3), narrowing the gap with the A states by about two NAEP scale score points.

The NAEP changes generally contradict the MSU ratings.

Changes in NAEP Scores (in scale score points), Group B States (13 states)

Table 3-4

MSU Rating	2009–2011	2011–2013	2009–2013
5 (n=7)	1.70	0.74	2.44
4 (n=1)	–0.72	3.68	2.96
3 (n=5)	3.08	0.14	3.22
All (n=13)	2.05	0.73	2.78

Source: Generated from NAEP Data Explorer data. MSU Rating obtained from William H. Schmidt's May 2012 presentation to the National Press Club, Washington, D.C. Categorical labels 1–5 assigned by author.

The progress states make in implementing the CCSS is crucial to the standards' impact on achievement.

This may indicate regression to the mean. There could also be another, unknown factor driving the relatively larger gains by Group B states. Whatever the cause, the gains in Group B states cast doubt on Schmidt and Houang's hypothesis that implementation difficulties lie at the heart of Group B's underperformance on the 2009 NAEP. If these states had trouble implementing their own standards prior to 2009, it is difficult to imagine them suddenly discovering the secret to implementation in the first few years with the Common Core. And if resource constraints were a primary factor hobbling past efforts at implementation, surely finding adequate resources during the Great Recession limited what the Group B states could accomplish.

In sum, the Schmidt and Houang ratings of state math standards in 2009 do not predict gains on NAEP very well in subsequent years. The notion that disaggregating the states into two groups would clarify matters because 13 states (Group B) faced implementation challenges also does not receive support. Whether in Group A or Group B, states with 2009 math standards most dissimilar to the Common Core made the largest NAEP gains from 2009–2013.

NAEP Changes in Light of Common Core's Implementation

As Schmidt and Houang point out—and any informed observer would surely agree—the progress states make in implementing the CCSS is crucial to the standards' impact on achievement. The MSU congruence ratings were designed to serve as substitutes for CCSS implementation measures, there being no implementation to measure in 2009. Now, with the passage of time, it is possible to get an early reading on implementation from a direct measure of state efforts. A 2011 survey of state educational agencies was mentioned above. The survey was conducted as part of a U.S. Department of Education study of reforms promoted by the Recovery Act. The Common Core was one such reform. The survey asked states if they had: (1) adopted the CCSS; (2) provided, guided, or funded professional development on the CCSS; (3) provided curriculum/instructional materials for the CCSS; and (4) worked with a consortium to develop assessments aligned with the CCSS.

For the current study, the states' responses were utilized to create an implementation rating. Modifications to the survey answers were made if a press report was located updating a state's status after the 2011 survey was conducted. Montana, Washington, and Wyoming, for example, had not yet adopted the CCSS when the survey was conducted, but they did soon thereafter. Georgia, Kansas, Oklahoma, Pennsylvania, and Utah have either withdrawn from their respective CCSS assessment consortium or announced a freeze on CCSS testing.

The category "non-adopter" was assigned to states that answered "no" to all four questions. That group consists of Alaska, Minnesota, Nebraska, Texas, and

Virginia. Those states are going their own way on math standards and can serve as a control group for CCSS.[39] At the other end of the implementation continuum, the category "strong" was assigned to states answering "yes" to all four questions. A total of 19 states have adopted the CCSS, taken steps to provide both professional development and curriculum/instructional materials aligned with CCSS, and are members of a consortium designing CCSS assessments. They are the strong implementers of CCSS. The remaining 26 states are medium implementers. They adopted CCSS but have not taken all of the other steps available to them to implement the standards.

Table 3-5 shows the average NAEP gains of the states based on implementation of CCSS. The 2009–2013 gains are what CCSS advocates hope for, at least in terms of a consistent pattern. The strong implementers made the largest gains (1.88), followed by the medium implementers (1.00), and then the non-adopters (0.61). The 2011–2013 pattern is also favorable towards the implementers of the CCSS. The medium implementers made the most progress (0.61) and the strong implementers made gains (0.21), although less than in 2009–2011. Caution must be exercised with the non-adopters since they only include five states, and Alaska's decline of 1.49 scale score points from 2009–2013 diminishes what was an average gain of more than one point by the other four states.

Changes in NAEP Scores (in scale score points), By Implementation of CCSS

Table 3-5

Implementation Rating	2009–2011	2011–2013	2009–2013
Strong (n=19)	1.65	0.23	1.88
Medium (n=26)	0.39	0.61	1.00
Non-adopters (n=5)	1.30	–0.69	0.61
All (n=50)	0.96	0.34	1.30

Note: Strong = adopted CCSS in math and pursued three implementation strategies (professional development, new instructional materials, joined testing consortium). Medium = adopted CCSS math standards but did not employ at least one of the implementation strategies. Non-adopters = did not adopt CCSS.

Source: Modified from data in Table H.1. "Standards and Assessment Indicators by State, 2010–2011," State Implementation of Reforms Promoted Under the Recovery Act, A. Weber, et al (2014).

Discussion

The Schmidt and Houang state standards ratings of 2009 proved to be a poor predictor of progress on NAEP in subsequent years. A rating based on states' implementation activities did reveal a pattern. States that more aggressively implemented the CCSS registered larger gains from 2009–2013. That's an optimistic finding for CCSS.

Let's evaluate the magnitude of potential gains from CCSS using that optimistic finding. Start by recognizing that from 1990–2013—the entire history of the main NAEP assessment—scores on the eighth grade math test rose from 263 to 285, a gain of 22 points. That averages to about one scale score point per year. The gains from 2009–2013 have significantly lagged that pace. As reported in Table 3-5, the average gain for the entire period was 1.30, which comes out to 0.33 per year. Critics of CCSS might suspect that the transition to CCSS is responsible for the slowing, but the data presented here do not support the charge. The five states that rejected the CCSS have performed worse than the states that adopted CCSS.

But how much worse? What is the difference? Not much. The 1.27 gap between strong implementers and non-adopters is about .035 of the 2009 NAEP's standard deviation (36). A rule of thumb is that differences of less than .20 SD are not even noticeable, let alone significant. If it takes four years for the CCSS to generate a .035 SD improvement, it will take 24 years for a noticeable improvement to unfold. And that

States that more aggressively implemented the CCSS registered larger gains from 2009–2013.

Even the most ardent Common Core supporter won't be satisfied with a three point NAEP gain.

improvement would add up to 7.62 NAEP scale score points, a gain in 24 years that falls far short of the 22 point gain that NAEP registered in its first 23 years.

Recent NAEP gains might be disappointing because the economic turmoil of the past few years presented an inopportune time for implementing new standards. That's possible, but the historical record is mixed. The early 1990s recession was accompanied by weak NAEP gains, but the early 2000s recession took place while NAEP scores were soaring. Perhaps the positive effects of the CCSS will not fully emerge until assessments aligned with the standards are administered and accountability systems tied to the results are launched. There is evidence that the test-based accountability systems of the late 1990s and the NCLB inspired systems of the early 2000s had a positive impact on achievement; however, in many jurisdictions, accountability systems were then being implemented for the first time.[40] The new CCSS accountability systems will be replacing systems that are already in place. The quality that they add to or subtract from existing systems is unknown. Moreover, as the consequences of NCLB's accountability systems began to be felt, significant political opposition arose in many states. Whether the CCSS systems experience the same backlash remains to be seen.

Can the small, insignificant effect of implementation be reconciled with the MSU study? Schmidt and Houang reported the tests of statistical significance for their congruence rating but they did not report an estimate of CCSS effects on NAEP scores. It is always possible for a statistically significant regression coefficient to denote an effect that is insignificant in the real world. Statistical significance tells us that we can be confident that an effect is different from zero, not that the difference is important. This is an especially relevant distinction when an analysis of NAEP data is conducted with states as the unit of analysis. As pointed out in a 2012 Brown Center Report study of the CCSS, most variation on NAEP lies within states—between students, not between states.[41] The standard deviation of state NAEP scores on the 2009 math test is 7.6 points. The standard deviation of the 2009 NAEP eighth grade math score, a statistic based on variation in student performance, is 36 points—four to five times larger.

An illustration of what these two SDs mean for interpreting the magnitude of CCSS effects is revealing. Schmidt and Houang's congruence rating has a range of 662–826, mean of 762, and SD of 33.5. The regression coefficient for the congruence rating was 0.08.[42] A statistical convention is to calculate the impact that a one SD change in an independent variable (in this case, the congruence rating) has on the dependent variable (in this case, the 2009 eighth grade NAEP score). In plain English, how much of a boost in NAEP scores can we expect from a pretty big increase in the congruence rating? A little arithmetic produces the following: a one SD gain in the congruence rating (33.5 points) is predicted to yield a NAEP gain of 2.68 points. Consider that gain in terms of the two SDs. It is about 0.35 of the state-level SD—a moderate but noticeable effect that is consistent with MSU's finding of statistical significance. But as a proportion of the student-level SD, the effect is only 0.07 SD, which is quite small, even undetectable. Moreover, the MSU analysis could not assign a firm estimate of how much time it took for states with standards similar to CCSS to generate this tiny effect, although six to eight years is a good guess.[43]

The point here is not that Schmidt and Houang did anything wrong. State level policies certainly can be evaluated with

state-level data. The problem is that a statistically significant finding from an analysis of state-level NAEP scores, the variation among states being relatively small, often fades to insignificance when considered in the more practical, real world terms of how much math students are learning. It is doubtful that even the most ardent Common Core supporter will be satisfied if the best CCSS can offer—after all of the debate, the costs in tax revenue, and blood, sweat, and tears going into implementation—is a three point NAEP gain.

The 2012 Brown Center Report predicted, based on an empirical analysis of the effects of state standards, that the CCSS will have little to no impact on student achievement. Supporters of the Common Core argue that strong, effective implementation of the standards will sweep away such skepticism by producing lasting, significant gains in student learning. So far, at least—and it is admittedly the early innings of a long ballgame—there are no signs of such an impressive accomplishment.

NOTES

1 Winnie Hu, "New Recruit in Homework Revolt: The Principal," *New York Times,* June 15, 2011, page a1.

2 Tom Loveless, "PISA's China Problem," The Brown Center Chalkboard, October 9, 2013, http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2013/10/09-pisa-china-problem-loveless; Tom Loveless, "Attention OECD-PISA: Your Silence on China is Wrong," The Brown Center Chalkboard, December 11, 2013, http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2013/12/11-shanghai-pisa-scores-wrong-loveless; Tom Loveless, "PISA's China Problem Continues: A Response to Schleicher, Zhang, and Tucker," The Brown Center Chalkboard, January 8, 2014, http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2014/01/08-shanghai-pisa-loveless.

3 Kam Wing Chan, *Ming Pao Daily News*, January 3, 2014, http://faculty.washington.edu/kwchan/ShanghaiPISA.jpg.

4 Li Tao, et al. *They are Also Parents: A Study on Migrant Workers with Left-behind Children in China* (Beijing: Center for Child Rights and Corporate Social Responsibility, August 2013).

5 Dexter Roberts, "Chinese Education: The Truth Behind the Boasts," *Bloomberg Businessweek*, April 4, 2013, http://www.businessweek.com/articles/2013-04-04/chinese-education-the-truth-behind-the-boasts.

6 Statistic comes from the most recent population estimate from the UAE National Bureau of Statistics, reported in 2010 and available at http://www.uaestatistics.gov.ae/EnglishHome/ReportDetailsEnglish/tabid/121/Default.aspx?ItemId=1914&PTID=104&MenuId=1.

7 Kam Wing Chan, op. cit.

8 Kam Wing Chan, *Ming Pao Daily News*, January 3, 2014, http://faculty.washington.edu/kwchan/ShanghaiPISA.jpg.

9 Andreas Schleicher, "Are the Chinese cheating in PISA or are we cheating ourselves?" *OECD Education Today,* December 10, 2013, http://oecdeducationtoday.blogspot.com/2013/12/are-chinese-cheating-in-pisa-or-are-we.html.

10 *OECD Economic Surveys: China 2013* (OECD Publishing, March 2013): pp. 91–92.

11 Anthony Jackson, "Shanghai Responds to School Ranking Cheating Allegations," Global Learning Blog, *Education Week,* December 21, 2013, http://blogs.edweek.org/edweek/global_learning/2013/12/shanghai_responds_to_school_ranking_cheating_allegations.html.

12 Marc Tucker and Andreas Schleicher, "Response to Brookings Institution Attack on PISA," Top Performers Blog, *Education Week*, December 26, 2013, http://blogs.edweek.org/edweek/top_performers/2013/12/response_to_the_brookings_institution_attack_on_pisa.html.

13 World Bank, "Population ages 0-14 (% of total)," World DataBank: World Development Indicators database, http://data.worldbank.org/indicator/SP.POP.0014.TO.ZS.

14 Schleicher, op. cit.

15 Marc Tucker, "Tom Loveless on Hukou in China," Top Performers Blog, *Education Week*, January 24, 2014, http://blogs.edweek.org/edweek/top_performers/2014/01/tom_loveless_on_hukou_in_china.html.

16 "Lessons from PISA for the United States," *Strong Performers and Successful Reformers in Education* (OECD Publishing, 2011): pp. 96.

17 Marc Tucker (Ed.), *Chinese Lessons: Shanghai's Rise to the Top of the PISA League Tables* (Washington, DC: National Center on Education and the Economy, 2014), pp. 2.

18 John Stevens, *Getting the Word Out About the National Assessment: The National Assessment Governing Board and Changes in NAEP Reporting* (National Assessment Governing Board. March 2009): pp. 7.

19 *PISA 2012 Results: What makes School Successful? Resources, Policies, and Practices* (Volume IV), (OECD Publishing, 2013).

20 Ibid.

21 Chris Cook, "Shanghai tops global state school rankings," *Financial Times*, December 7, 2010, http://www.ft.com/intl/cms/s/0/20770bf6-01e7-11e0-b66c-00144feabdc0.html; C. M. Rubin, "The Global Search for Education: Can You Pass the Global Standardized Test?" *The Huffington Post*, May 17, 2011, http://www.huffingtonpost.com/c-m-rubin/global-standardized-testing_b_862165.html; Sean Coughlan, "China: The world's cleverest country?" BBC News, May 12, 2012, http://www.bbc.co.uk/news/business-17585201.

22 Cook, op. cit.

23 Jim Stevenson (Narr.), *Daybreak Asia*, December 10, 2013, http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2014/01/08-shanghai-pisa-loveless.

24 Sean Coughlan, "China: The world's cleverest country?" *BBC News*, May 12, 2012, http://www.bbc.co.uk/news/business-17585201.

25 Brian Gill and Steven Schlossman, "A Sin Against Childhood: Progressive Education and the Crusade to Abolish Homework, 1897-1941," *American Journal of Education*, vol. 105, no. 1 (Nov., 1996), 27–66. Also see Brian P. Gill and Steven L. Schlossman, "Villain or Savior? The American Discourse on Homework, 1850-2003," *Theory into Practice*, 43, 3 (Summer 2004), pp. 174–181.

26 Sara Bennett and Nancy Kalish, *The Case Against Homework: How Homework is Hurting Our Children and What We Can Do About It,* (New York: Crown, 2006). John Buell, *Closing the Book on Homework: Enhancing Public Education and Freeing Family Time,* (Philadelphia: Temple University Press, 2004). Alfie Kohn, *The Homework Myth: Why Our Kids Get Too Much of a Bad Thing,* (Cambridge, MA: Da Capo Press, 2006). Etta Kralovec and John Buell, *The End of Homework: How Homework Disrupts Families, Overburdens Children, and Limits Learning,* (Boston: Beacon Press, 2000).

27 Hu, op. cit.

28 Data for other years are available on the NAEP Data Explorer. For Table 1, the starting point of 1984 was chosen because it is the first year all three ages were asked the homework question. The two most recent dates (2012 and 2008) were chosen to show recent changes, and the two years in the 1990s to show developments during that decade.

29 NAEP's sampling design lessens the probability of skewing the homework figure. Students are randomly drawn from a school population, meaning that an entire class is not tested. Teachers would have to either single out NAEP students for special homework treatment or change their established homework routine for the whole class just to shelter NAEP participants from homework. Sampling designs that draw entact classrooms for testing (such as TIMSS) would be more vulnerable to this effect. Moreover, students in middle and high school usually have several different teachers during the day, meaning that prior knowledge of a particular student's participation in NAEP would probably be limited to one or two teachers.

30 NAEP Question B003801 for 9 year olds and B003901 for 13- and 17-year olds.

31 Met Life, *Met Life Survey of the American Teacher: The Homework Experience*, November 13, 2007, pp. 21–22.

32 Karl Taro Greenfeld, "My Daughter's Homework Is Killing Me," *Atlantic*, September 18, 2013.

33 William H. Schmidt, *New Research Links Common Core Math Standards to Higher Achievement*. Presented at the National Press Club, May 3, 2012. PowerPoint available at http://www.achieve.org/files/BILL_CCSSM.ppt.

34 William H. Schmidt and Richard T. Houang, "Curricular Coherence and the Common Core State Standards for Mathematics," *Educational Researcher* 41, no. 8 (2012): 294–308.

35 William H. Schmidt, Curtis C. McKnight, Gilbert A. Valverde, Richard T. Houang, and David E. Wiley, *Many visions, many aims, Volume I: A crossnational investigation of curricular intentions in school mathematics* (Dordrecht: Kluwer Academic Publishers, 1997); William H. Schmidt, Curtis C. McKnight, Senta A. Raizen, Pamela M. Jakwerth, Gilbert A. Valverde, Richard G. Wolfe, Edward D. Britton, Leonard J. Bianchi, and Richard T. Houang, *A splintered vision: An investigation of US science and mathematics education*. Vol. 3. (Boston: Kluwer Academic Publishers, 1997); William H. Schmidt and Richard T. Houang, "Lack of Focus in the Mathematics Curriculum: Symptom or Cause?" in *Lessons Learned: What International Assessments Tell Us about Math Achievement*, ed. Tom Loveless (Washington: Brookings Institution Press, 2007).

36 Schmidt and Houang, op. cit., p. 307.

37 Jan-Eric Gustafsson, "Understanding Causal Influences on Educational Achievement through Analysis of Differences over Time within Countries," in *Lessons Learned: What International Assessments Tell Us about Math Achievement*, ed. Tom Loveless (Washington: Brookings Institution Press, 2007).

38 The 2011 survey of state education agencies is reported by Webber et. al, "State Implementation of Reforms Promoted Under the Recovery Act," U.S. Department of Education (January 2014).

39 Minnesota adopted the CCSS in English-Language Arts, but not mathematics.

40 For a description of standards and accountability systems in the late 1990s, see Education Week, *Quality Counts: Rewarding Results, Punishing Failure* (January 11, 1999). For evidence of positive effects

of test based accountability, see Thomas Dee and Brian A. Jacob, "Evaluating NCLB," *Education Next* 10, no. 3 (Summer 2010); Manyee Wong, Thomas D. Cook, and Peter M. Steiner, "No Child Left Behind: An Interim Evaluation of Its Effects on Learning Using Two Interrupted Time Series Each with Its Own Non-Equivalent Comparison Series," Working Paper 09-11 (Evanston, IL: Northwestern University Institute for Policy Research, 2009); Eric A. Hanushek and Margaret E. Raymond, "Does school accountability lead to improved student performance?" *Journal of Policy Analysis and Management* 24, no.2 (Spring 2005): 297–327; Martin Carnoy and Susanna Loeb, "Does external accountability affect student outcomes? A cross-state analysis," *Educational Evaluation and Policy Analysis* 24, no. 4 (2002): 305–331.

41 Tom Loveless, *The 2012 Brown Center Report on American Education* (Washington: The Brookings Institution, 2012).

42 Summary statistics for rating of congruence is on page 300 and regression output is on page 304 of William H. Schmidt and Richard T. Houang, "Curricular Coherence and the Common Core State Standards for Mathematics," *Educational Researcher* 41, no. 8 (2012): 294–308.

43 The No Child Left Behind (NCLB) Act of 2001required standards in all states. About half of the states already had standards in place before NCLB. Schmidt and Houang's estimate is that eighth graders in 2009 probably spent all or most of their school years learning math under the state standards then in place.

THE BROOKINGS INSTITUTION

STROBE TALBOTT
President

DARRELL WEST
Vice President and Director
Governance Studies Program

BROWN CENTER STAFF

GROVER "RUSS" WHITEHURST
Senior Fellow and Director

TOM LOVELESS
Senior Fellow

MATTHEW CHINGOS
Fellow

BETH AKERS
Fellow

KATHARINE LINDQUIST
Research Analyst

SARAH WHITFIELD
Center Coordinator

LIZ SABLICH
Communications Manager

Views expressed in this report are solely those of the author.

BROWN CENTER on
Education Policy
at BROOKINGS

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

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