NEPC

NATIONAL EDUCATION POLICY CENTER

# Review of *Math Performance in Global Perspective*


*Reviewed By*

Jeremy Kilpatrick

University of Georgia

January 2011


## Summary of Review


A report from Harvard's Program on Education Policy and Governance and the journal *Education Next* finds that only 6% of U.S. students in the high school graduating class of 2009 achieved at an advanced level in mathematics compared with 28% of Taiwanese students and more than 20% of students in Hong Kong, Korea, and Finland. Overall, the United States ranked behind most of its industrialized competitors. The report compares the mathematics performance of high achievers not only across countries but also across the 50 U.S. states and 10 urban districts. Most states and cities ranked closer to developing countries than to developed countries. However, the study has three noteworthy limitations: (a) internationally, students were sampled by age and not by grade, and countries varied greatly on the proportion of the student cohort included in the compared grades; in fact, only about 70% of the U.S. sample would have been in the graduating class of 2009, which makes the comparisons unreliable; (b) the misleading practice of reporting rankings of groups of high-achieving students hides the clustering of scores, inaccurately exaggerates small differences, and increases the possibility of error in measuring differences; and (c) the different tests used in the study measured different domains of mathematics proficiency, and the international measure was limited because of relatively few test items. The study's deceptive comparison of high achievers on one test with high achievers on another says nothing useful about the class of 2009 and offers essentially no assistance to U.S. educators seeking to improve students' performance in mathematics.

**National Education Policy Center**

School of Education, University of Colorado
Boulder, CO 80309-0249
Telephone: 303-735-5290
Fax: 303-492-7090

Email: NEPC@colorado.edu
http://nepc.colorado.edu

Publishing Director: Alex Molnar

GREAT LAKES CENTER
FOR EDUCATION RESEARCH & PRACTICE

This is one of a series of Think Twice think tank reviews made possible in part by funding from the Great Lakes Center for Education Research and Practice. It is also available at http://greatlakescenter.org.

# Review of *Math Performance in Global Perspective*

*Jeremy Kilpatrick, University of Georgia*

## I. Introduction

A new study sponsored by the Program on Education Policy and Governance at Harvard University and the journal *Education Next* examines the prevalence of high-achieving students domestically and internationally. The study[1] compared the percentages of U.S. students in the 50 states and 10 urban districts who performed at an advanced level in mathematics on the 2005 National Assessment of Educational Progress (NAEP) with estimated percentages of students in other countries who would have reached that same level had they taken the NAEP 2005 mathematics assessment. In the case of NAEP mathematics, *advanced level* means the student scored at or above a score arbitrarily set by the National Assessment Governing Board. That so-called cut-score has long been criticized as "too high,"[2] and the report itself acknowledges that critics have called it "excessively stringent"[3] while going on to say "we do not take a position on this question" (p. 10). Only 6% of U.S. eighth graders reached the advanced level on NAEP 2005 for mathematics.

Samples of U.S. students in the graduating class of 2009 had taken NAEP 2005 when they were eighth graders and the Program for International Student Assessment (PISA) 2006 when they were tenth graders, so a statistical procedure could be used to calculate the PISA mathematics literacy score that was presumed equivalent to scoring at an advanced level on NAEP. It was thus possible to estimate what percentage of students in each of the 57 countries in PISA 2006 would have attained that advanced NAEP level. The report indicates that most of those countries had far greater percentages of high achievers than the United States did. The results varied by state and by urban district, but the percentages in most, in the words of the report, were "shockingly below those of many of the world's leading industrialized nations" (p. 14).

The report attempts to refute two possible explanations for the low percentage of high achievers in the United States: (a) a relatively large number of U.S. students from immigrant families and historically marginalized groups, and (b) the No Child Left Behind (NCLB) legislation's purported shifting of resources from high to low achievers. Data are presented in the report to suggest that neither explanation is plausible.

In previous studies, U.S. eighth graders' average performance in mathematics compared favorably with that of eighth graders in other countries. The new report argues that those findings were obtained because relatively few of those other countries were members of the Organization for Economic Co-operation and Development (OECD), on whose members' performance the PISA international average is based.[4]

As indicated in this review, there are six items of concern that readers of the report should consider:

1. NAEP is taken by eighth graders, whereas PISA is taken by 15-year-olds in each country, who are typically not all going to be in the same graduating class.
2. The mathematical proficiency of those students who graduated in the class of 2009 was likely rather different from what it had been in 2005 or 2006.
3. To rank districts and states measured on one scale against countries measured on another gives a misleading picture of how they are performing, especially when the scores come from one tail of a distribution (the very highest achievers) and are estimated using statistical linking.
4. Although a statistical linkage had in the past been possible between NAEP and previous international tests that had been administered to equivalent representative samples of students in mathematics at the same time and in the same grades, "no similar basis exists to create a linkage between PISA and NAEP."[5]
5. NAEP mathematics is designed to measure the knowledge, skills, and competencies needed by U.S. students in mathematics at different grades, whereas PISA is not intended to be tied to the school mathematics curriculum, calling instead on the ability to use and apply knowledge and skills in real-world situations; the study thus tries to mesh the results of two different tests that measure different domains of mathematics proficiency.
6. Mathematics was a so-called minor domain in PISA 2006, which means that there was likely substantial variability within countries across the test content, raising unresolved reliability and validity concerns.[6]

## II. Findings and Conclusions of the Report

The study found that, of the 57 countries participating in PISA 2006, 30 had a higher percentage of high achievers in mathematics than the United States when their PISA scores were used to estimate how the countries would have done on the NAEP 2005 mathematics assessment. For example, 28% of the students in Taiwan were high achievers in mathematics, and more than 20% of the students in Hong Kong, Korea, and Finland reached that level—compared with only 6% in the United States.

When the states and urban districts were ranked along with the PISA countries according to their percentage of high achievers, the state with the highest percentage, Massachusetts, had only 11% of its students in the category of high achievers, and 14 countries had higher percentages. The lowest ranking states—West Virginia, New Mexico, and Mississippi—had lower percentages of high-performing students than did 43 countries, including Serbia and Uruguay. At the district level, although Austin, Charlotte, and Boston had higher percentages than the U.S. as a whole, the remaining cities—such as Washington, Chicago, Los Angeles, and Atlanta—had lower percentages, typically at the level of developing countries in the PISA sample.

To address the argument that the low percentage of high-achieving students in the U.S. might well stem from its heterogeneous population, with large numbers of students from immigrant families or historically marginalized groups, the report includes percentage estimates at the advanced level for white students and for children of parents with a college degree. Only 8% of the U.S. students who reported their ethnicity as white and only 10.3% of those who reported that at least one parent had a college degree reached the advanced level. The percentages were higher than these for all students, regardless of their ethnicity or parents' educational attainment, in 24 and 16 other PISA countries, respectively.

The report cites a different report, from the Fordham Institute and authored by Tom Loveless,[7] to support the contention that the 2002 No Child Left Behind legislation may have shifted attention and resources to lower-performing students and thus taken resources away from students at the advanced level. That, if it were true, might have had the effect of decreasing their top-end test scores. Data in the report show, however, that over the years from 1996 to 2009, the percentage of eighth-grade students at the advanced level in NAEP mathematics has steadily increased, and the percentage below the basic level has steadily decreased.

The report cites two earlier studies by Gary Phillips[8] comparing the average performance in NAEP 2007 mathematics of eighth graders in each of the 50 U.S. states with that of eighth-grade students in other countries as assessed by the Trends in International Mathematics and Science Study (TIMSS) in 2003 and 2007, respectively. In the earlier studies, the results were much more favorable to the U.S. students, but those studies focused on average performance while the new study focuses on just the highest-achieving students. The authors of the new report do not, however, attribute the different findings to the focus of the earlier studies on average performance. Instead, they argue that the differences arose because of differences between PISA and TIMSS in the set of participating countries. The set of countries in TIMSS 2003 and 2007 was more heavily weighted with low-performing developing countries than was PISA 2006.

The report concludes by suggesting that although students' performance in mathematics is shaped by many factors within and outside schools, "some of our findings point specifically to problematic elements within the nation's schools" (p. 24). It claims that because so few relatively advantaged white students and students with a college-educated parent were high achievers in mathematics, U.S. "schools are failing to teach students effectively" (p. 24). The authors do not identify those "problematic elements," nor do they explain how the low percentages of high achievers in mathematics stem from ineffective teaching rather than other factors.

Given that the United States already has a high rate of expenditure per pupil in Grades K–12, the authors contend that more money is not likely to help. Proposing no policy changes that might lead to better mathematics performance, the authors of the report end by noting that although NCLB appears not to be harming high-performing students, its laudable policy emphasis on low-performing students may be inadvertently taking attention away from high-performing students. Because of confusion as to how *high performing* is defined, however, the report's conclusions are unsound, and its utility minimal.

## III. The Report's Rationale for Its Findings and Conclusions

The report argues that "a cadre of highly talented professionals trained to the highest level of accomplishment is needed to foster innovation and growth" (p. 6) in the United States. Today, however, "the performance of U.S. schools, once the envy of the world, has slipped" (p. 6), and the demand for high achievers in mathematics and science is not being met. Attention is being given, due to NCLB, to basic levels of achievement, especially among disadvantaged students, but there is a great need for more students to reach high levels of achievement.

Mathematics is the focus of interest in the study because it "appears to be the subject in which accomplishment in secondary school is particularly significant for both an individual's and a country's economic well-being" (p. 8).[9] Furthermore, "there is a fairly clear international consensus on the math concepts and techniques that need to be mastered and on the order in which those concepts should be introduced into the curriculum" (p. 9), something that is not true of reading or science.

NAEP mathematics and PISA mathematics were chosen as the vehicles for examining how high achievers in the United States compared with high achievers in other countries because NAEP 2005 had been given to a representative sample of U.S. eighth graders between January and March 2005,[10] and PISA 2006 was administered between September and November 2006, when most of those same students were in tenth grade.[11]

In an article entitled "Your Child Left Behind" in the December 2010 *Atlantic* magazine, Amanda Ripley pointed out why the report's principal author Eric Hanushek chose to analyze the performance of white students and students having at least one parent with a college education:

> These days, the theory Hanushek hears most often is what we might call the diversity excuse. When he runs into his neighbors at Palo Alto coffee shops, they lament the condition of public schools overall, but are quick to exempt the schools their own kids attend. "In the litany of excuses, one explanation is always, 'We're a very heterogeneous society—all these immigrants are dragging us down. But our kids are doing fine,'" Hanushek says. This latest study was designed, in part, to test the diversity excuse.[12]

As indicated above, the report's findings and conclusions rule out the so-called diversity excuse and demonstrate that across the nation the percentage of U.S. students considered high achieving in mathematics is much smaller than the corresponding percentage in other countries.

## IV. The Report's Use of Research Literature

The report draws appropriately on much of the extensive literature on NAEP, PISA, and TIMSS as well as the literature on economic consequences of academic skill development and on what the authors see as the growing need for investment in human capital. Some of that literature consists of earlier contributions by the report's authors. Somewhat ironically, however, in view of the choice of mathematics proficiency as a focus, almost no serious attention is given to the literature on differences in the mathematics assessed across the three assessments. The authors observe in one footnote, "Some have suggested that PISA and NAEP do not test a common domain of knowledge" (p. 10). They go on to cite Mark Schneider's argument that "PISA is a self-proclaimed 'yield study' assessing the 'literacy' of 15-year-olds and is not tied to any specific curricula"[13] and Tom Loveless's critique of the validity of the PISA science test.[14] But then they turn around and argue that "average student performances across countries on different international tests are strongly correlated, suggesting that a common domain of knowledge is being tested" (p. 10). As noted below, the correlation coefficient they cite does not imply that the domain is common.

The authors ignore several efforts[15] to consider differences among the three assessments in the mathematics being measured. In a view across the content differences in science and mathematics among the three assessments, a document from the National Center for Education Statistics concluded: "*The mathematics and science being assessed differ in terms of the ways in which the frameworks for assessment are organized and in terms of content coverage, item format, and other key features*."[16] Analyses not considered in the report indicate that the features of NAEP and TIMSS are much more similar than those of PISA. For example, PISA differs considerably from TIMSS and NAEP in item type. "Only about a third of the PISA items are multiple choice, in contrast to roughly two thirds of the TIMSS items"[17] and three quarters of the NAEP items.[18] Moreover, relatively more PISA items concern data analysis and probability, whereas relatively more TIMSS and NAEP items concern algebra.[19] One analysis summarized the difference between the TIMSS and PISA mathematics assessments as follows: "Where TIMSS questions are more direct and abstract, PISA questions are more lengthy and wordy. TIMSS asks, 'What do you know?' while PISA asks, 'What can you do?'"[20] Again, these analyses are not considered in the new report; instead, the concern is dismissed based on strong between-country correlations between different assessments. The further apart two assessments are in what they are measuring, however, the riskier is any attempt to equate them statistically, particularly when one is considering only the upper tails of the distributions.

## V. Review of the Report's Methods

The authors of the report used what Phillips terms *statistical moderation*[21] to find the score on PISA 2006 mathematics that was equivalent to scoring at the advanced level on NAEP 2005 mathematics. As they note, "A score on PISA 2006 of 617.1 points is equivalent to the lowest score obtained by anyone in the top 6.04 percent of U.S. students in the Class of 2009" (p. 10). They then found the percentage of students in each country who scored at or above 617.1 on PISA and ranked those percentages along with the percentages of students in each urban district and state who scored at the advanced level on NAEP. The resulting graph for countries and states, Figure 1 (p. 16), has 95 bars (50 states and 45 countries; 12 countries were excluded from the graph because they had less than 1% of their students at the advanced level) that range from high to low. Two similar graphs (Figures 2 & 3, pp. 17–18) are included that have the same bars for the countries but that have bars giving the percentage of white students in each U.S. state at the advanced level and the percentage of students with at least a college-educated parent in each state at the advanced level, respectively. A separate bar graph (Figure 4, p. 19) ranks the 10 urban districts and 45 countries from high to low on the percentage of students at the advanced level.

To link the NAEP and PISA scores, the authors needed to make several questionable assumptions; namely, "that both NAEP and PISA tests randomly select questions from a common universe of mathematics knowledge" (p. 10) and "that students who scored similarly on the two exams will have similar math knowledge" (p. 11). As noted in the above discussion of the relevant research literature, several studies call those assumptions into doubt. The questions on the two tests are, according to experts who study those assessments, taken from substantially different universes of knowledge, meaning that similar scores do not necessarily reflect similar mathematics proficiency.

## VI. Review of the Validity of the Findings and Conclusions

### Class of 2009

The authors of the report say, "We focus on the performance of the international equivalent of the U.S. high school [graduating] Class of 2009 at the time when this class was in the equivalent of U.S. grades 8 and 9" (p. 9). The term *Class of 2009* is never defined, nor is *international equivalent,* but the usage is problematic.

First, the set of U.S. students who graduated from high school in 2009 was different from the set of students who were in the eighth grade in the spring of 2005 and even more different from the 15-year-olds who were in the tenth grade in the fall of 2006. Moreover, of U.S. 15-year-olds in the PISA 2006 sample, only between two-thirds and three-fourths were in Grade 10,[22] and the percentage who graduated in 2009 was undoubtedly less than that. So one cannot know what the percentage of high achievers might then have been. Further, the unknown comparability of dropout rates across the PISA countries from Grade 8 to 12 makes any conclusions about the class of 2009 speculative at best.

Second, by the time the class of 2009 graduated, the students had, depending on the country, state, district, and program they were in, studied mathematical content in their last two years of secondary school that was not adequately measured in NAEP or PISA. For example, although "PISA items classified to a higher level of cognitive complexity and demand than NAEP items, the content covered was most consistent with the topics specified in the NAEP eighth-grade framework"[23] and not with the twelfth-grade framework. Further, PISA does not address knowledge of advanced mathematics topics, and NAEP does not assess advanced topics such as trigonometry, conic sections, vectors, normal distributions, or conditional probability until Grade 12.[24] Consequently, by 2009 the set of high achievers might have changed significantly.

Third, some members of the class of 2009 or its international equivalent were well assessed by PISA 2006 but others not at all. Table 1 shows, for five selected countries in PISA, the distribution of the percentages of 15-year-olds at each grade. Finland had no students in Grade 10, whereas Japan had all of them there. Defining a cohort by grade, as in NAEP, is clearly different internationally from defining it by age, as in PISA.

**Table 1: Percentage Distribution of 15-Year-Old Students by Grade Level for Selected PISA 2006 Countries**

| Country | Grade 7 | Grade 8 | Grade 9 | Grade 10 | Grade 11 | Grade 12 |
|---|---|---|---|---|---|---|
| *USA* | 1 | 1 | 11 | 71 | 17 | 0 |
| *Finland* | 0 | 12 | 88 | 0 | 0 | 0 |
| *Germany* | 2 | 12 | 55 | 28 | 0 | 0 |
| *Japan* | 0 | 0 | 0 | 100 | 0 | 0 |
| *UK* | 0 | 0 | 0 | 1 | 98 | 1 |

Source: National Center for Education Statistics Report NCES 2008–016, Table C-1.

## Advanced levels

Students from other countries who earned a PISA mathematics score at or beyond the 93.96th percentile for U.S. students were certainly doing well on that assessment, *but they might or might not have performed equally well had they taken the NAEP assessment some 18 months earlier*. In 2006, science literacy was the major domain in PISA, and mathematics literacy was a minor domain. It was assessed with only 48 items,[25] which is not much of base on which to establish the particular percentage of students in a given country who were high performers in mathematics. The PISA 2006 mathematics item set had to address the entire range of performance, and relatively few items were used to identify the high scorers. In contrast, NAEP "is designed to also provide estimates for individual states, which requires an increased sample size; and thus measures performance at a higher level of precision than . . . PISA."[26]

The authors of the new report essentially conceded the imprecision of restricting attention to one tail of a distribution when, in discussing their lack of emphasis on advanced-level performance across countries in science and reading, they argued that setting a standard so high that only 3% of U.S. students could attain it "isolates such a small percentage of the population that it introduces considerable error in measuring cross-country differences. Simply put, the tests may be subject to considerable error when viewed at a cutoff so far from the average performance" (p. 26). Although somewhat attenuated, the same criticism applies when the standard is set at 6%. The authors report in appendices the standard errors for the percentages they calculate, but the bar graphs they use in the text to rank those percentages obscure the fact that many adjacent percentages are not significantly different. When standard errors are accounted for, it becomes clear how meaningless it is to provide a ranked list of countries and states; no difference is found between a state like Minnesota ranked 20th in the list and a country like Sweden ranked as 24th.

## Mathematics proficiency

NAEP mathematics items are designed to reflect "common curricular practices in the nation's schools and ultimately are intended to reflect the best thinking about the knowledge, skills, and competencies needed for students to have a deep level of understanding at different grades."[27] In contrast, PISA focuses on mathematics literacy, which includes the mastery of processes, understanding of concepts, and application of knowledge in various situations, drawing "not only from school curricula but also from learning that may occur outside of school."[28]

The authors rest their claim that PISA and NAEP assess a common domain of knowledge in large part on evidence that "the correlation between average student performances across countries on the PISA 2006 and TIMSS 2007 is 0.93" (p. 35), a point they make twice.[29] That correlation is calculated for the 27 countries that participated in both assessments. The implication is that PISA and NAEP would be equally highly correlated given that NAEP and TIMSS "measure essentially the same mathematical content"[30] and can be easily linked.

Just because two instruments are highly correlated, however, it does not follow that one can replace the other. For example, in a study of TIMSS 1995, Erling Boe and his colleagues found that across 40 countries, the knowledge and skills required of eighth graders to persist in completing, and being able to complete, the more than 100-item TIMSS background questionnaire (as measured by the national average percentage of items answered) account for 60% of the skills and knowledge required to complete the mathematics achievement test (as measured by the national mean achievement score).[31] The 60% means that the correlation coefficient is about .78. By the logic of replacing one variable by its correlate, persistence in answering a questionnaire about one's home environment, school experiences, and attitudes could be taken as a measure of one's mathematics proficiency. As another example, the correlations at the country level in PISA between the domains of mathematical, scientific, and reading literacy are all above .90.[32] But that correlation does not mean at all that those three domains can be interchanged. Not only does correlation not imply causality; it does not imply equivalence.

## VII. Usefulness of the Report for Guidance of Policy and Practice

The claim made in the report that relatively few students in the United States score at an advanced level in mathematics, whether measured by NAEP, PISA, or TIMSS, has been well established. What is misleading is that the percentages of advanced-level students in countries, states, and districts can be put on the same scale so that they can easily be compared. The bar graph on pages 16 of the report indicates, for example, that in mathematics, the percentage of advanced-level students in the class of 2009 in Finland far exceeded that of the percentage in Massachusetts, which in turn exceeded that of the percentage in the United Kingdom. Yet almost none of either the Finnish or the U.K. students in PISA would have been in the class of 2009 (see Table 1, above).

The term *advanced level* has two meanings in this study. The first refers to the arbitrary cut-score on NAEP 2005 mathematics that U.S. eighth graders must reach or exceed to be deemed advanced. Percentages of eighth graders in states and urban districts who are "advanced" can thus be calculated and ranked. The second meaning refers to a corresponding cut-score on PISA 2006 mathematical literacy that 15-year-olds in PISA countries must reach or exceed to be deemed advanced. Percentages of 15-year-olds in those countries who are "advanced" by that measure can also be calculated and ranked, although those rankings are subject to greater error than the NAEP rankings because of the NAEP test's greater precision. So far, so good.

But the correspondence between the two cut-scores rests on using the same percentile (93.96th) of U.S. eighth graders in 2005 and U.S. 15-year-olds in 2006. Putting the percentages of countries, states, and urban districts on the same bar graph amounts to endorsing two scenarios. Scenario 1 says, “If eighth graders in PISA countries had taken the NAEP mathematics assessment in 2005, then these are the percentages that would have scored at the advanced NAEP level.” Scenario 2 says, “If 15-year-olds in U.S. states and urban districts had taken the PISA mathematical literacy assessment in 2006, then these are the percentages that would have scored at the advanced PISA level.” Neither of the scenarios’ hypotheses is in fact true, but the authors of the report would like the reader to accept Scenario 2. Further obscuring the statistical slight of hand, the authors label the bar graph with the term “class of 2009,” even though none of the data were obtained from twelfth graders in 2009.

At the website of the National Center for Education Statistics, state and urban district policymakers can obtain ample data on how their students at all levels of proficiency have been performing in NAEP mathematics. Those policymakers do not need, and should avoid, this flawed effort to enter their high-performing graduating seniors of 2009 into a mock international horse race in which they did not participate. It is just as well that the report offers no recommendations for “policy changes that might foster excellence” (p. 24). Any such recommendations might have been as untenable as the study itself.

## Notes and References

[1] Hanushek, E.A., Peterson, P.E., & Woessmann, L. (2010, November). *U.S. math performance in global perspective: How well does each state do at producing high-achieving students?* (PEPG Report No. 10–19). Cambridge, MA: Harvard's Program on Education Policy and Governance and *Education Next*. Retrieved December 2, 2010, from
http://www.hks.harvard.edu/pepg/PDF/Papers/PEPG10-19_HanushekPetersonWoessmann.pdf.

[2] According to a 1999 report from the National Research Council, "Numerous external comparison studies conducted by the National Academy of Education supported the conclusion that NAEP cutscores between the proficient and advanced levels and between the basic and proficient levels are consistently set too high, with the outcome of achievement-level results that do not appear to be reasonable relative to numerous other external comparisons." See p. 167 of

Pellegrino, J.W., Jones, L.R., & Mitchell, K.J. (Eds.). (1999*). Grading the nation's report card: Evaluating NAEP and transforming the assessment of educational progress*. Committee on the Evaluation of National and State Assessments of Educational Progress, Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education, National Research Council. Washington, DC: National Academy Press.

[3] Loveless, T. (2008). Part I: Analysis of NAEP data. In A. Duffen, S. Farkas, & T. Loveless (Eds.*), High-achieving students in the era of No Child Left Behind*. Washington, DC: Thomas B. Fordham Institute.

[4] In addition to the mathematics discussion, appendices of the report present comparable data for science and reading performance (science for the class of 2009 but reading for the class of 2006), details of study methods, and discussion of the sets of countries used in previous international comparative studies.

[5] Phillips, G.W. (2009, June). *The second derivative: International benchmarks in mathematics for U.S. states and school districts*. Washington, DC: American Institutes for Research, 4. Retrieved December 20, 2010, from
http://www.air.org/files/International_Benchmarks1.pdf.

[6] Mazzeo, J., & von Davier, M. (2008). *Review of the Programme for International Student Assessment (PISA) test design: Recommendations for fostering stability in assessment results* (Education Working Papers EDU/PISA/GB(2008)28). Paris: OECD, 23-24. Retrieved December 27, 2010, from
https://edsurveys.rti.org/PISA/documents/MazzeoPISA_Test_DesignReview_6_1_09.pdf.

[7] Loveless, T. (2008). Part I: Analysis of NAEP data. In A. Duffen, S. Farkas, & T. Loveless (Eds.*), High-achieving students in the era of No Child Left Behind*. Washington, DC: Thomas B. Fordham Institute.

[8] Phillips, G.W. (2007, November). *Chance favors the prepared mind: Mathematics and science indicators for comparing states and nations*. Washington, DC: American Institutes for Research. Retrieved December 20, 2010, from
http://www.air.org/files/phillips.chance.favors.the.prepared.mind.pdf.

Phillips, G.W. (2009, June). *The second derivative: International benchmarks in mathematics for U.S. states and school districts*. Washington, DC: American Institutes for Research. Retrieved December 20, 2010, from
http://www.air.org/files/International_Benchmarks1.pdf.

[9] Two authors of the report demonstrated in an earlier study "that growth in the economic productivity of a nation is driven more clearly by the math proficiency of its high school students than by their proficiency in other subjects" (p. 9). See Table 2 in

Hanushek, E.A., & Woessmann, L. (2009, January). *Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation* (NBER Working Paper 14633). Cambridge, MA: National Bureau of Economic Research.

[10] See *NAEP 2005 assessment results – Information for educators: FAQs*. Retrieved January 10, 2011, from
http://nationsreportcard.gov/2005_assessment/s0041.asp?printver=.

[11] The report claims that most of the U.S. 15-year-olds who took part in PISA 2006 were in Grade 9 (p. 9), but most were actually in Grade 10 when the test was administered in the fall. See pp. 4 and 42 of

Baldi, S., Jin, Y., Skemer, M., Green, P.J., & Herget, D. (2007). *Highlights from PISA 2006: Performance of U.S. 15-year-old students in science and mathematics literacy in an international context* (NCES 2008–016). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

[12] Ripley, A. (2010, December). Your child left behind. *Atlantic*. Retrieved December 20, 2010, from http://www.theatlantic.com/magazine/archive/2010/12/your-child-left-behind/8310/.

[13] Schneider, M. (2009, Fall). The international PISA test: A risky investment for states. *Education Next, 9*(4), 68–75.

[14] Loveless, T. (2009, January). *How well are American students learning?* (2008 Brown Center Report on American Education, Vol. 2, No. 3). Washington, DC: Brookings Institution.

[15] National Center for Education Statistics. (2007). *Comparing TIMSS with NAEP and PISA in mathematics and science*. Washington, DC: Author. Retrieved December 23, 2010, from http://www.eric.ed.gov/PDFS/ED503624.pdf.

Grønmo, L.S., & Olsen, R.V. (2007). TIMSS versus PISA: The case of pure and applied mathematics. In *The Second IEA International Research Conference: Proceedings of the IRC-2006, Vol. 1: Trends in International Mathematics and Science Study (TIMSS)* (pp. 201–214). Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement.

Neidorf, T.S., Binkley, M., Gattis, K., & Nohara, D. (2006). *Comparing mathematics content in the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and Program for International Student Assessment (PISA) 2003 assessments* (NCES 2006-029). Washington, DC: National Center for Education Statistics. Retrieved December 23, 2010,

from
http://nces.ed.gov/pubs2006/2006029.pdf.

Wu, M. (2009). A comparison of PISA and TIMSS 2003 achievement results in mathematics and science. *Prospects, 39*(1), 33–46. doi: 10.1007/s11125-009-9109-y Retrieved December 23, 2010, from http://www.springerlink.com/content/5438748557v01g71/.

[16] National Center for Education Statistics. (2007). *Comparing TIMSS with NAEP and PISA in mathematics and science*. Washington, DC: Author, 6 (italics in original). Retrieved December 23, 2010, from
http://www.eric.ed.gov/PDFS/ED503624.pdf.

[17] Kilpatrick, J. (2009, February-March). TIMSS 2007 mathematics: Where are we? *MAA Focus: The Newsmagazine of the Mathematical Association of America, 29*(7), 4–7, 6.

[18] National Center for Education Statistics. (2007). *Comparing TIMSS with NAEP and PISA in mathematics and science*. Washington, DC: Author, 12. Retrieved December 23, 2010, from http://www.eric.ed.gov/PDFS/ED503624.pdf.

[19] National Center for Education Statistics. (2007). *Comparing TIMSS with NAEP and PISA in mathematics and science*. Washington, DC: Author, Table 3. Retrieved December 23, 2010, from http://www.eric.ed.gov/PDFS/ED503624.pdf.

[20] Hutchison, D., & Schagen, I. (2007). Comparisons between PISA and TIMSS: Are we the man with two watches? In T. Loveless (Ed.), *Lessons learned: What international assessments tell us about math achievement* (pp. 227–261). Washington, DC: Brookings Institution Press. 254.

[21] Phillips, G.W. (2007, November). *Chance favors the prepared mind: Mathematics and science indicators for comparing states and nations*. Washington, DC: American Institutes for Research, 91. Retrieved December 20, 2010, from
http://www.air.org/files/phillips.chance.favors.the.prepared.mind.pdf.

[22] Baldi, S., Jin, Y., Skemer, M., Green, P.J., & Herget, D. (2007). *Highlights from PISA 2006: Performance of U.S. 15-year-old students in science and mathematics literacy in an international context* (NCES 2008–016). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. P. 4.

[23] Stephens, M., & Coleman, M. (2007). *Comparing PIRLS and PISA with NAEP in reading, mathematics, and science* (Working Paper). U.S. Department of Education. Washington, DC: National Center for Education Statistics, 11. Retrieved January 10, 2011, from
http://nces.ed.gov/Surveys/PISA/pdf/comppaper12082004.pdf.

[24] Neidorf, T.S., Binkley, M., Gattis, K., & Nohara, D. (2006). *Comparing mathematics content in the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and Program for International Student Assessment (PISA) 2003 assessments* (NCES 2006-029). Washington, DC: National Center for Education Statistics, Appendix A. Retrieved December

23, 2010, from
http://nces.ed.gov/pubs2006/2006029.pdf.

[25] Organization for Economic Co-operation and Development. (n.d.). *Comparisons over time on the PISA scales*. Paris: Author, 8. Retrieved December 20, 2010, from
http://www.oecd.org/dataoecd/19/34/39712449.pdf.

[26] Stephens, M., & Coleman, M. (2007). *Comparing PIRLS and PISA with NAEP in reading, mathematics, and science* (Working Paper). U.S. Department of Education. Washington, DC: National Center for Education Statistics, 12. Retrieved January 10, 2011, from
http://nces.ed.gov/Surveys/PISA/pdf/comppaper12082004.pdf.

[27] Neidorf, T.S., Binkley, M., Gattis, K., & Nohara, D. (2006). *Comparing mathematics content in the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and Program for International Student Assessment (PISA) 2003 assessments* (NCES 2006-029). Washington, DC: National Center for Education Statistics, 5. Retrieved December 23, 2010, from
http://nces.ed.gov/pubs2006/2006029.pdf.

[28] National Center for Education Statistics. (2007). *Comparing TIMSS with NAEP and PISA in mathematics and science*. Washington, DC: Author, 2. Retrieved December 23, 2010, from
http://www.eric.ed.gov/PDFS/ED503624.pdf.

[29] Each time they cite

Phillips, G.W. (2009, June). *The second derivative: International benchmarks in mathematics for U.S. states and school districts*. Washington, DC: American Institutes for Research, 36. Retrieved December 20, 2010, from
http://www.air.org/files/International_Benchmarks1.pdf.

[30] Phillips, G.W. (2009, June). *The second derivative: International benchmarks in mathematics for U.S. states and school districts*. Washington, DC: American Institutes for Research, 10. Retrieved December 20, 2010, from
http://www.air.org/files/International_Benchmarks1.pdf.

[31] Boe, E.E., May, H., & Boruch, R.F. (2002). *Student task persistence in the Third International Mathematics and Science Study: A major source of achievement differences at the national, classroom, and student levels* (Research Rep. No. 2002-TIMSS1). Philadelphia, PA: University of Pennsylvania, Graduate School of Education, Center for Research and Evaluation in Social Policy.

[32] Organization for Economic Co-operation and Development. (2009). *PISA data analysis manual: SPSS* (2nd ed.). Paris: Author, 26. Retrieved December 20, 2010, from
http://www.oecd.org/document/31/0,3746,en_2649_35845621_42592351_1_1_1_1,00.html.

DOCUMENT REVIEWED: **U.S. Math Performance in Global Perspective: How Well Does Each State Do at Producing High-Achieving Students?**

AUTHORS: Eric A. Hanushek, Paul E. Peterson, Ludger Woessmann

PUBLISHERS: Harvard's Program on Education Policy and Governance & Education Next

DOCUMENT RELEASE DATE: November 9, 2010

REVIEW DATE: January 11, 2011

REVIEWER: Jeremy Kilpatrick, University of Georgia

E-MAIL ADDRESS: jkilpat@uga.edu

PHONE NUMBER: (706) 542-4163

SUGGESTED CITATION:

Kilpatrick, J. (2011). *Review of "U.S. Math Performance in Global Perspective: How Well Does Each State Do at Producing High-Achieving Students?"* Boulder, CO: National Education Policy Center. Retrieved [date] from http://nepc.colorado.edu/thinktank/review-us-math.