

What are PISA and TIMSS doing in *Science*?

Brad Wible
Senior Editor

National Academy of Education
16 Sep 2016



Science

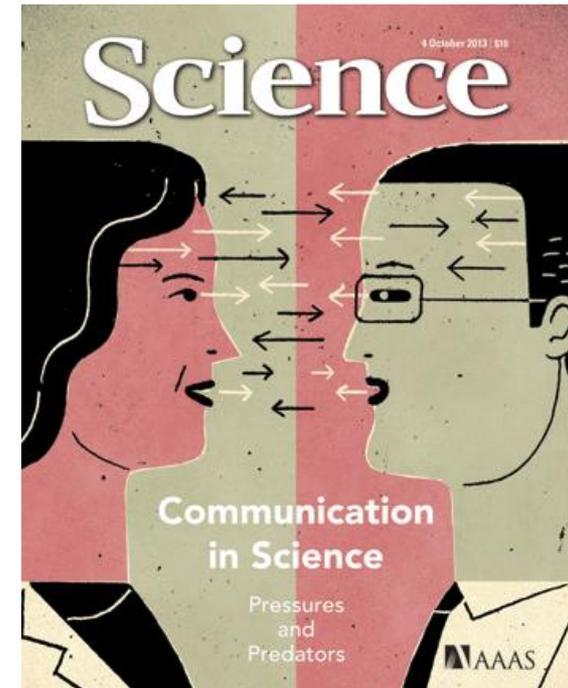
AAAS

Not unique to ILSAs

Communication, public understanding

Integrating science and policy

“Hype” of research results



EDUCATION FORUM | SCIENCE EDUCATION

Climate confusion among U.S. teachers

Eric Plutzer¹, Mark McCaffrey², A. Lee Hannah³, Joshua Rosenau², Minda Berbeco², Ann H. Reid²

+ Author Affiliations

*Corresponding author. E-mail: plutzer@psu.edu

Science 12 Feb 2016:

What is AAAS ?

- Founded in 1848
- World's largest general science society
- 120,000 members from 91 countries
- Annual meeting hosts 7,000+ attendees
- S&T policy & mass media fellows
- Many more education, advocacy, diplomacy efforts

print circulation of ~130k

~6M/month unique visitors
online

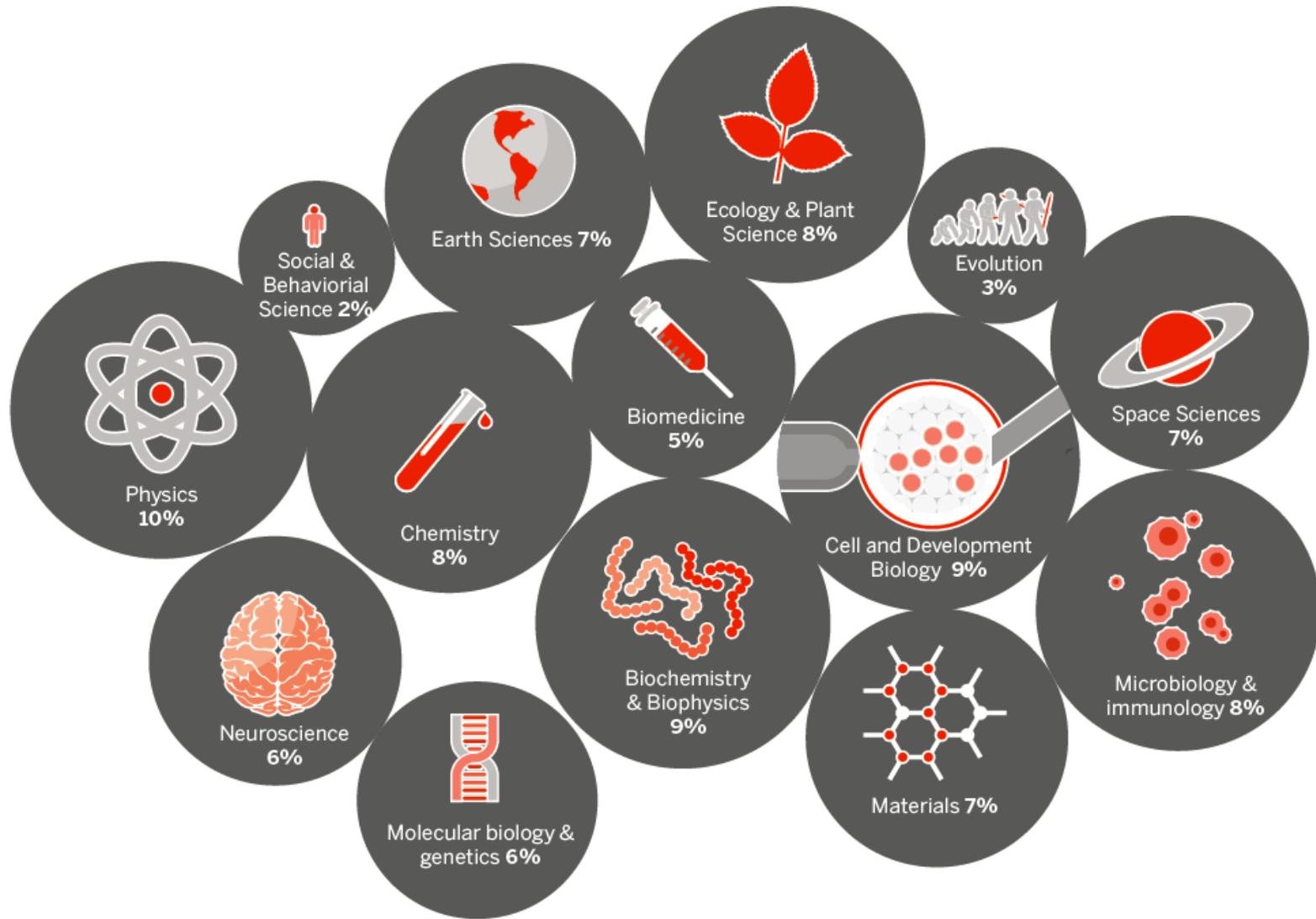
Over 80% of papers were
in the 95th percentile of
altmetrics



675K followers



2.8M fans



Who we are, what we do

Authors and readers more technically oriented than most, though still written for broader audience

Do we cover education, ILSAs? Yes

But the cover says **Science**, not **Education**

Focus on: STEM education

Scientific study of education

Nobody focused solely on education, so coverage competes with many (all) other topics, every week

Who we are, what we do

Research

Editorials

News

Policy Forum, Education Forum

Book Reviews

Editor's Choice / In Other Journals

Literature Reviews

Special Issues: Education and Technology, Literacy,
Early Childhood Education, Grand Challenges



SCIENCE EDUCATION

Will latest U.S. law lead to successful schools in STEM?

Despite shortcomings, there is room for cautious optimism

By Adam Gamoran^{1,2*}

Remarkably, after a long delay, the U.S. Congress passed and President Obama signed in late 2015 a new version of the Elementary and Secondary Education Act, the law that since 1965 has staked out the federal government's commitment to education from kindergarten to 12th grade (K-12). This version, the Every Student Succeeds Act (ESSA), replaces the controversial version known as No Child Left Behind (NCLB) (1, 2). With ESSA scheduled to take full effect in fall 2017, states are now in planning mode. What changes are in store for schooling in science, technology, engineering, and mathematics (STEM)? Although Congress sought advice on this topic, few changes in ESSA reflect new knowledge on STEM. Yet the law provides opportunities for improvement if states choose to act.

velop its own accountability system by setting standards and assessments, as in NCLB, but also to determine what counts as "success" and to establish its own incentives and supports for improvement. Whereas NCLB accountability focused solely on reading and math, ESSA is open to "multiple measures" of achievement, including those for science, if so desired by states.

During the NCLB era, Congress was often frustrated with academic achievement across U.S. schools. The House Committee on Appropriations, Subcommittee on Commerce, Justice, Science, and Related Agencies—which oversees the National Science Foundation (NSF) budget—requested that NSF provide a report identifying characteristics of effective schools and programs in K-12 STEM (3). The result was a report by the Board on Science Education of the National Research Council (NRC), now known as the National Acade-

How did NRC, at Congress' request, advise on best evidence for effective K-12 STEM education?

Does 2015 Every Student Succeeds Act reflect this advice?

Reason to be optimistic for K-12 STEM education upon implementing ESSA in 2017?

ILSA in Science: Editorials

28 Jan 2005

Math and Science Achievement

Rodger W. Bybee, executive director of the Biological Science Curriculum Study

Donald Kennedy, Editor-in-Chief of Science

The 2003 TIMSS results contain some important messages about what's happening globally in math and science education ... We think that a strong argument can be made for linking national math and science standards to the improvement we see in the U.S. TIMSS results. ... TIMSS is always a sobering test for U.S. educators, because the results force a comparison of their system—which insists on the right of 50 states and 15,000 school districts to shape their own mathematics and science programs—with centrally organized systems such as that of Singapore, the world leader. The results represented by TIMSS 2003 would suggest giving accelerated emphasis to strong voluntary national standards. ... Perhaps the 2003 TIMSS results will, in addition to providing a baseline for comparing international educational systems, persuade U.S. policy-makers to count the costs of leaving everything to local decision-making.

29 Aug 2008

Academies Active in Education

Jorge E. Allende, IAP Science Education Program, former president, Chilean Academy of Sciences.

The results of PISA reveal that all developing countries and many industrial ones, including the United States, are failing to prepare their children adequately for life in the modern world.

28 Aug 2009

Great Teachers for STEM

Maxine Singer, president emeritus, Carnegie Institution of Washington

... in the 2006 PISA, American 15-year-olds ranked 24th out of 57 countries in science and 32nd in mathematics.

ILSA in Science: Editorials

23 Oct 2009

Europe Rethinks Education

Pierre Léna, Délégué à l'éducation in the French Académie des Sciences.

... Yet the performance of young people from France and Germany on PISA is barely average, and many students leave middle school being illiterate in science.

16 Apr 2010

Common U.S. Math Standards

Philip Daro, Director of the Strategic Education Research Partnership

William McCallum, University of Arizona

Jason Zimba, Bennington College

... studies such as TIMSS have shown that U.S. mathematics performance is compromised by a lack of focus and coherence in the curriculum. Hong Kong students outscore U.S. students on the grade 4 TIMSS assessment, even though the Hong Kong curriculum covers about half of the tested topics, whereas the U.S. grade 4 curriculum covers over 80% of the tested topics. Higher-performing countries teach central topics more coherently and in greater depth. The proposed new standards differ substantially from today's typical state standards by focusing on key topics and building a coherent progression of learning.

28 May 2010

Seeking Science Standards

Alan I. Leshner, Chief Executive Officer of AAAS and Executive Publisher of Science

Shirley Malcom, Director of AAAS Education and Human Resources Programs

Jo Ellen Roseman, Director of AAAS Project 2061

... Nearly all of America's competitor countries have national science education standards and score much higher on international science achievement assessments: U.S. 15-year-olds ranked 21st among students in 30 developed nations in science on the 2006 PISA.

ILSA in Science: Editorials

13 Jan 2012

It's the Teachers

John E. Burris, President of the Burroughs Wellcome Fund

The first Trends in International Mathematics and Science Study (TIMSS) results were released in 1995 and the first Program for International Student Assessment (PISA) test results in 2000. The education reforms that they helped to motivate have had little impact on U.S. performance, and the country continues to hope for a simple solution that will miraculously turn the tide. Finland has scored near the top of the PISA examinations for the past decade, and the lessons of its success are simple: Recruit the best and the brightest to be teachers, and train them extensively and well.

27 Jan 2012

Teaching Real Science

Bruce Alberts, Editor-in-Chief of Science

In this issue of Science, we are publishing the first of 15 winning entries for the 2011 Science Prize for Inquiry-Based Instruction ... Created for introductory college science courses, each module can be readily used in many different settings and schools. ... We hope that these contests will help support a rethinking of science education that is consistent with the new Framework for K-12 Science Education (precollege) from the U.S. National Academies, as well as with one of the central goals in the international PISA Science Competencies: "Understands the characteristic features of science as a form of human knowledge and enquiry.

ILSA in Science: News

10 Dec 2004

Hong Kong, Finland Students Top High School Test of Applied Skills

David Grimm

PISA, which covers science, math, and reading literacy, complements a set of tests called the TIMSS, which measures fourth- and eighth-graders' knowledge of specific concepts, such as geometrical formulas and chemical principles. PISA takes the premise a step further by measuring how students apply the sum of this education to new problems. "We're not asking whether students can read," says Thomas Romberg, a math educator at the University of Wisconsin, Madison, who helped design a version of the PISA exam administered in 2000 that focused primarily on reading literacy but included science and math questions. "We're seeing whether they can understand a book they've never seen before." ... As for PISA's impact on U.S. science and math education, Suter also believes that state assessments should be reevaluated to gauge the application of knowledge, not just retention, as a marker of student progress.

28 Sep 2007

U.S. Says No to Next Global Test of Advanced Math, Science Students

Jeffrey Mervis

The NCES says it is bowing out of 2008 TIMSSA, an advanced version of the TIMSS given quadrennially to younger students, because it can't fit the \$5 million to \$10 million price tag into its flat budget. Officials also question whether the target cohort—students finishing secondary school who have taken advanced mathematics and physics courses—is comparable around the world.

ILSA in Science: News

10 Dec 2010

Shanghai Students Lead Global Results on PISA

Jeff Mervis

A group of teenagers from Shanghai, China, have posted the top scores on the latest version of an international test of practical knowledge in reading, mathematics, and science. It's the first time that students from mainland China have participated in PISA.

8 Jul 2011

Korean Students Are Top Digital Readers

South Korean teenagers beat out their peers from 18 other countries in the first assessment of digital literacy. ... One surprising finding: Greater use of the computer at home didn't necessarily translate into a better score. ... The test is part of a plan to expand PISA beyond pencil-and-paper assessments; future assessments will include computer-based tests of problem-solving skills.

21 Dec 2012

East Asian Students Still Excel in Math, Science

"Revolutionary results require revolutionary changes," says Michael Martin, co-director of the International Study Center at Boston College that administers the quadrennial TIMSS. Those changes are more likely to occur, he says, in countries that have a centralized education system and can move quickly to embrace the latest thinking on how to improve schooling.

3 Dec 2013

Asia Remains on Top in Test of Student Knowledge

Jeffrey Mervis

In addition to compiling head-to-head rankings, OECD collects data on educational policies and practices around the world, with topics ranging from how countries educate their poorest students to parental attitudes toward the importance of an education.

ILSA in Science: Education Forum

25 May 2007

Cognitive Supports for Analogies in the Mathematics Classroom

Lindsey E. Richland, Osnat Zur, Keith J. Holyoak

We investigated how certain mathematics classroom activities differ between the United States and nations in which students score higher on international tests. We focused on factors of cognition and memory, which can be distinguished from cultural differences in instruction. The video portion of the TIMSS, a large-scale international video study of classroom mathematics instruction, indicated that American teachers introduced conceptually connected, rich problems at rates similar to teachers from higher-achieving countries. However, they engaged students in complex connected reasoning and problem-solving substantially less often.

30 May 2008

Culture, Gender, and Math

Luigi Guiso, Ferdinando Monte, Paola Sapienza, Luigi Zingales

... To assess the relative importance of biological and cultural explanations, we studied gender differences in test performance across countries. Cultural inequalities range widely across countries, whereas results from cognitive tests do not. We used data from the 2003 PISA ... The tests were designed by the OECD to be free of cultural biases. ... To explore the cultural inputs to these results, we classified countries according to several measures of gender equality. (i) The World Economic Forum's Gender Gap Index reflects economic and political opportunities, education, and well-being for women. (ii) From the World Values Surveys, we constructed an index of cultural attitudes toward women based on the average level of disagreement to such statements as: "When jobs are scarce, men should have more right to a job than women." ...

ILSA in Science: Education Forum

28 Nov 2008

Global Sex Differences in Test Score Variability

Stephen Machin, Tuomas Pekkarinen

Do boys and girls differ in their intellectual and cognitive abilities and, if so, in what way? These questions have raised considerable debate, both in terms of average performance and in terms of variability around the average. ... We examined boy-girl differences in test score performance from the 2003 PISA, a survey of 15-year-olds who are enrolled in full- or part-time education in industrialized countries. Here, we analyze mathematics and reading test scores by country, focusing on differences in the mean and variance of these test scores.

2 Jan 2009

Technology and Testing

Edys S. Quellmalz, James W. Pellegrino

A new generation of technology-enabled assessments offers the potential for transforming what, how, when, where, and why testing occurs. ... Online testing now occurs in numerous international, national, and state assessment programs. The 2009 PISA will include electronic texts to test reading. ... Science assessment is perhaps leading the way in exploring the presentation and interpretation of complex, multifaceted problem types and assessment approaches. In 2006, PISA pilot-tested a Computer-Based Assessment of Science specifically to test knowledge and inquiry processes not assessed in the paper-based booklets.

ILSA in Science: Education Forum

23 Apr 2010

Assessing Literacy Across a Changing World

Andreas Schleicher

... Panels of international experts establish PISA literacy frameworks through three interrelated dimensions: (i) knowledge or structure of knowledge that students need to acquire; (ii) competencies that students need to apply; and (iii) contexts in which students would encounter problems. For example, students at the highest PISA level of literacy (level 5) must be able to locate and use information that is difficult to find in unfamiliar texts, to show detailed understanding and infer which information is relevant to the task, to evaluate critically, to build hypotheses, to draw on specialized knowledge, and to accommodate concepts that may be contrary to expectations. By contrast, students at the baseline level of proficiency (level 2) only need to be capable of basic reading tasks, such as locating straightforward information, making low-level inferences of various types, working out what a well-defined part of a text means, and using outside knowledge to understand it.

10 Jun 2011

Preparing Future Math Teachers

William H. Schmidt, Richard Houang, Leland S. Cogan

... What is the role for recruitment versus what must be accomplished through preparation? ... We reexamine data from the 2010 Teacher Education and Development Study in Mathematics (TEDS-M), a 16-country survey of math teachers-in-training near the end of their final semester. TEDS was conducted by the IEA with an eye to developing international benchmarks for teacher preparation. This is similar to what was done for K-12 curricula via the IEA TIMSS. ... Another way to explore this issue is to compare the knowledge of potential future middle school teachers of mathematics, as reflected by TEDS, with the average eighth-grade mathematics achievement, as found in TIMSS. The strong relation between TEDS and TIMSS [$R^2 = 0.70$] reflects how selection and/or recruitment based on knowledge before entry into teacher training relates to future teachers' knowledge when exiting from training.

ILSA in Science: Education Forum

23 Sep 2011

The Pseudoscience of Single-Sex Schooling

Diane F. Halpern, et al

In attempting to improve schools, it is critical to remember that not all reforms lead to meaningful gains for students. We argue that one change in particular—sex-segregated education—is deeply misguided, and often justified by weak, cherry-picked, or misconstrued scientific claims rather than by valid scientific evidence. ... Large-scale reviews in Great Britain, Canada, Australia, and New Zealand, as well as analyses of data from the PISA, similarly found little overall difference between SS and mixed-sex academic outcomes.

29 Jun 2012

European Teacher Training Reforms

Johannes Bauer, Manfred Prenzel

... In many European countries, teacher education has recently undergone substantial transformations during the Bologna Process of higher-education reform. What trends these reforms have invoked and whether they are for better or for worse in terms of teacher education is unclear. ... Evaluating benefits of these developments for attaining better-educated teachers will require establishment of a continuous international monitoring system similar to TIMSS and PISA in secondary education.

ILSA in Science: Education Forum

22 Jan 2016

Knowledge capital, growth, and the East Asian miracle

Eric A. Hanushek, Ludger Woessmann

... Over the past quarter-century, both theoretical and empirical analyses of possible drivers of the different growth rates seen around the world invariably assign an important role to human capital. This has led to development policies focused on increasing enrollment and retention in schools. We argue, however, that too much attention is paid to the time spent in school, and too little is paid to the quality of the schools and the types of skills developed there. ... We combine all available results for each country from international math and science tests taken between 1964 and 2003 (extended with Latin American regional tests) into a single direct measure of human capital, which, in the aggregate, we call the knowledge capital of a nation. Our work provides an approach for aggregating scores across previously unintegrated tests. This combines skills developed in schools and also in families, among peers, and through cultures.

ILSA in Science: Book Reviews

13 Jul 2007

Can Evidence Inform the Debate?

Marcia C. Linn

Why Aren't More Women in Science? Top Researchers Debate the Evidence

Stephen J. Ceci and Wendy M. Williams, Eds.

... The 15 essays bring to life recent findings on the involvement of women and men in science courses and careers. ... The contributors present their strongest arguments, support those with their best data, and articulate their beliefs about the current participation of women in science ... several of the essays consider performance on competitive mathematics tests, such as the SAT and the TIMSS, often offering rather divergent interpretations of such evidence.

2 Jan 2009

What's Wrong with Inferences from Test Scores?

Edward Haertel

Measuring Up: What Educational Testing Really Tells Us

Daniel Koretz

By and large, people who hear reports that scores on state tests have increased since last year, that the black-white achievement gap is closing, or that 81% of Massachusetts fourth graders are proficient readers think they understand pretty well what those reports mean. The logic of using test scores to evaluate schools, inform educational decisions, and influence curriculum and instruction is rarely questioned. ... It turns out that there are a lot of problems and that we would do well to try and understand them better. ... Pointing out that even large score differences may show inconsistencies from one test to another, Koretz discusses the 2003 eighth-grade mathematics gap between the United States and Norway. That year, the TIMSS placed the U.S. students far ahead of their Norwegian counterparts, while the PISA placed Norway ahead of the U.S. by a statistically significant margin.

ILSA in Science: Book Reviews

15 Jun 2012

Let's Finnish Schooling

Henrik Saalbach

Finnish Lessons: What Can the World Learn from Educational Change in Finland?

Pasi Sahlberg

... One small country in northern Europe did not follow the global trends when reforming its educational system—and it has done best: Finland, the primus of the PISA studies. Pasi Sahlberg's *Finnish Lessons* provides an insightful account of how, within recent decades, the Finns managed to shift their educational system from mediocre to top-performing.

ILSA in Science: Editor's Choice

17 Jul 2009

Correlates of the Gender Gap

Nosek et al., PNAS 106, 10593 (2009)

... Nosek et al. have studied the relation between gender stereotyping in the general population and student performance on TIMSS. In their virtual laboratory, any visitor can take an implicit association test (IAT) in any of 17 languages. In more than 500,000 tests collected from 2000 to 2008, roughly 70% of participants tended to associate words representing male with science faster than with liberal arts, and associated words representing female with liberal arts faster than with science. Across 34 countries, the male-female gap measured in the TIMSS correlated with the association of science and male as assessed in the IAT, with one standard deviation in stereotyping equivalent to 6.3 points on the standardized tests. The association of implicit (but not explicit) stereotypes in adults with national test scores in kids suggests that initiatives aimed at reducing the gap will need to be multifaceted.

21 Dec 2012

Summing Up Math Standards

Schmidt and Houang, Educational Researcher 41, 294 (2012)

Are the Common Core State Standards for Mathematics (CCSSM), which are about to be implemented in the United States, high-quality standards that are internationally competitive? Using data from standards of the highest-achieving nations on the 1995 TIMSS and all 50 state standards in place in 2008, Schmidt and Houang analyzed the focus, coherence, and rigor of the content defined by the CCSSM in an effort to predict their impact on student achievement. Comparison of CCSSM with the international standards revealed an almost 90% degree of consistency, suggesting that the CCSSM are focused, rigorous, and worthy of being world-class standards. ... Implementing the CCSSM will be worth the effort, however, as further analysis revealed that states with existing standards most similar to the CCSSM had higher NAEP scores.

ILSA in Science: Reviews

27 May 2011

Dyscalculia: From Brain to Education

Brian Butterworth, Sashank Varma, Diana Laurillard

Developmental dyscalculia is a mathematical disorder, with an estimated prevalence of about 5 to 7%, which is roughly the same prevalence as developmental dyslexia. ... The OECD demonstrated that an improvement of “one-half standard deviation in mathematics and science performance at the individual level implies, by historical experience, an increase in annual growth rates of GDP per capita of 0.87%”. Time-lagged correlations show that improvements in educational performance contribute to increased GDP growth. A substantial long-term improvement in GDP growth (an added 0.68% per annum for all OECD countries) could be achieved just by raising the standard of the lowest-attaining students to the PISA minimum level. At Level 1, students can answer questions involving familiar contexts in which all relevant information is present and the questions are clearly defined. They are able to identify information and carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are obvious and follow immediately from the given stimuli.

19 Apr 2013

Proficiency in Science: Assessment Challenges and Opportunities

James W. Pellegrino

... This Review argues that much of what is needed to effectively assess science learning, either at the classroom level or for purposes of system monitoring, has yet to be created and that design and implementation challenges are substantial. Even so, there are promising cases from which to learn and build. ... Although the quality of state assessments varies, none approximates the performance expectations discussed in the NRC Framework and NGSS. In contrast, there are two large-scale assessment programs that more closely exemplify aspects of science proficiency that involve science practices: the U.S. NAEP and the PISA.

ILSA in *Science*: Lessons

Yes, a few simple “US ranks X on PISA”, without much further analysis; mostly on Editorial page

Note key features, innovations, limitations in ILSAs

Analysis of ILSA data beyond basic rankings

Integrate ILSA data with other international data, to answer the questions being asked

Feedback from colleagues

If things don't change much, how do I write a story? Harder to sell as an incremental update.

Do they change the baseline? Can we easily compare with past years and across countries?

Two point change doesn't really mean anything ... or it means a lot ... or ...

Piecemeal results, hodgepodge, but what's the overarching idea/story? Changes in 4th grade math, but not in science, changes in 8th grade reading, but not math.

Finland going down, are they getting worse, or others getting better? Did Germany get better because of something they saw in the last tests and took measures to improve?

Feedback from colleagues

So many variables for making international comparisons.

Researchers think it's important because they have a lot of good data, but have they said anything new or useful? What questions are they trying to answer?

If US is so bad in STEM, why is US a world leader in S&T in so many respects? Why is US still an education destination? What are PISA and TIMSS about, if not clearly quality of education? Train to do well on a test, or to do well in advanced S&T?

Can the research inform broad scale changes in how people and organizations operate?

NSF has never asked us to address any of this when seeking funding for education projects.

Feedback from colleagues

Journalists did not offer much feedback on ILSAs (did they find them to be useful data springboards, did they feel inclined to cover them, etc)

Science content on ILSAs has gotten minimal media pickup

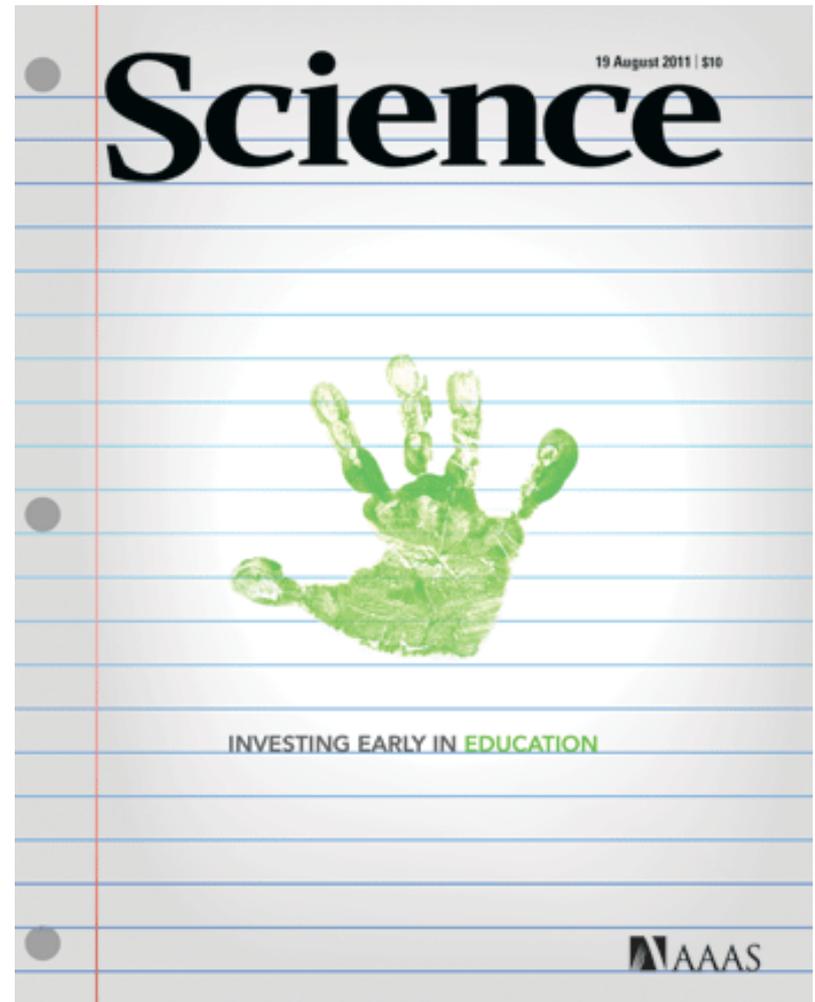
Work to find compelling (but not hyperbolic) ways to recognize their value to the press, e.g., fleshing out story possibilities for them based on newest ILSA kernels

Who is hungriest for these assessments?

Are country differences reflecting ability, or motivation? E.g., if US students had more reason to care, would they do better?

Thanks.

bwible@aaas.org



Science

