

PISA TESTING

A GLOBAL EDUCATIONAL RACE?

■ Svein Sjøberg – DOI: <https://doi.org/10.1051/ejn/2017402>

■ University of Oslo, ILS, Dept of Teacher Education and School Research Norway – svein.sjoberg@ils.uio.no

The OECD's PISA project is not an educational project. It is a political project and has to be understood as an instrument of power. PISA is normative, it tells what young people should learn, regardless of the nations' culture, nature, traditions and values. The battle to improve PISA rankings may conflict with our work to make science relevant, contextualized, interesting and motivating for young learners.



PISA, The Programme for International Student Assessment, was initiated by the OECD (Organization for Economic Cooperation and Development) in 2000 and has since then dramatically changed educational debates as well as policies worldwide. Although scientists and educators may appreciate the thinking behind the framework for testing, they also need to understand the wider social, ideological context of PISA.

The PISA project has to be understood as a social phenomenon and a political project, in essence a well-funded instrument of power that has steadily increased its influence on the educational discourse and policies in the now 70 participating countries. The educational debate has become global, and the race to improve PISA-rankings has become high priority in many countries.

While no feedback is given to students, teachers or schools, for governments the PISA-test is a high-stakes test. Governments are blamed for low scores, and governments are quick to take the honour when results are improving. Perceived bad rankings often create a crisis or panic, and governments are urged to do "something" to improve scores. But PISA cannot, by its design, say anything about cause and effect. Ill-informed school reforms are often introduced to improve PISA scores. National curricula, cultural values and priorities are pushed aside.

Many of the reforms that are legitimized by PISA can be characterized as New Public Management and neoliberal policies. Key words in these reforms are standardization of curricula, more testing, belief in competition, privatization and free choice of schools.

A global standard for all schools?

International studies are important, and can help us see our own country's education in perspective. Such studies, also of students' achievement in science and mathematics have been around more than 60 years. The acronym TIMSS (Trends in Mathematics and Science Study) is well known, and has provided interesting data for some decades. This study is descriptive and analytical, testing traditional knowledge and skills as close as possible to school curricula. TIMSS has been run mainly by academic institutions, and has no political or educational agenda. TIMSS researchers have by and large been careful in drawing causal conclusions and making recommendations, but TIMSS has influenced policy development as well as public debates in many countries.

The scene changed dramatically when the OECD entered the scene with its PISA-project in 2000. Since the first publication of results in 2001, the PISA-results have become a kind of global "gold standard" for educational quality - a single measure of the quality of the entire school system. An OECD-report on the policy impact of PISA, proudly states that "PISA has become accepted as a reliable instrument for benchmarking student performance worldwide, and PISA results have had an influence on policy reform in the majority of participating countries/economies" (Breakspear 2012).

Similarly, Andreas Schleicher (2012), director of PISA and recently also of Directorate of Education and Skills in OECD, in a TED talk starts his presentation by stating that PISA is "really a story of how international comparisons have globalized the field of education that we usually treat as an affair of domestic policy."

▲ ©iStockPhoto

What Does PISA Claim to Measure?

How does PISA present the project, what do they claim to measure? The following was stated already in 1999, before the first PISA testing took place in 2000:

How well are young adults prepared to meet the challenges of the future? Are they able to analyse, reason and communicate their ideas effectively? Do they have the capacity to continue learning throughout life? (OECD, 1999, p. 7)

If these points had been formulated as research questions in a research proposal for a 2 hour test for 15-year old students, the proposal would certainly have been rejected as too ambitious and unrealistic. Nevertheless, exact these words have been repeated in all later PISA reports.

In essence, the PISA creators are claiming that they have identified the essential skills and competencies necessary for future life, for all young learners, regardless of country, culture, history and nature.

So, although PISA explicitly states that it does not test school knowledge and that it does not test according to national curricula or test school knowledge, the PISA results are interpreted by OECD officials and policymakers around the globe as valid measures of the quality and the efficiency of national school systems.

The PISA test is anonymous, and no results are reported back to the students, their teachers or even schools. Results are only reported at the "system level", most often at the national level. When results are released in early December every third year, the main attention is given to the ranking of countries as given by the mean scores. The media reports are like reporting from sport events. If you google the term PISA and shock, scandal or crisis, you find hits from all over the globe.

The context of school science

In most countries, science is part of the school curriculum through the whole of its compulsory years. Scientists should always remember that the prime purpose of school science in these years is not to recruit the future scientist. Science at this level should contribute to the wider goals, ideals and values that the country has for its public school system, giving a foundation for future life and development for each individual. School science should open the eyes for young people to appreciate science and

to understand how important science is for our culture, our world-view and for meeting the challenges of the future. **While science as academic disciplines may be universal and independent of culture, school science needs to be embedded in the culture, and build on the interests and concerns of the learner.**

For the young learner, interest, enjoyment and engagement in science-related activities may be more important than achieving high test scores at this level. Positive attitudes are likely to last longer than the actual knowledge content that is taught. For some learners, positive experiences with school science may be the start of a science-related career, for others it might mean a lasting positive attitude to science. Test-driven teaching and rankings may be counterproductive, both for the possible future scientist and for the great majority.

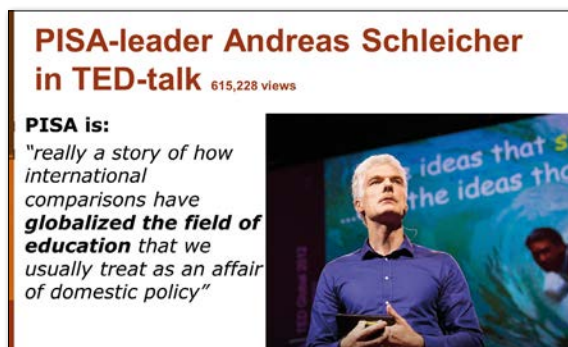
International tests need to be culture-neutral and decontextualized by design, if intended to measure an internationally valid construct. The selection process for PISA test items actually assures that no country or culture is favoured. Items that in the trial do not "function properly" are removed. So, while educators in science and other subjects argue for "context-based" teaching, tests like PISA have to avoid items that are contextualized in order to make a "fair test". Therefore, if higher rankings on international standardized tests become a *priority* for the national authorities, one may miss a key concern: to show the relevance of science and to stimulate the interest and curiosity in science.

The power of PISA

PISA is organized by the OECD, an organization that is "owned", governed and financed by its member states, the highly industrialized countries of the world. The prime mandate of the OECD is to promote economic growth and development in a global free market economy. Over the last decades, the importance of human resources and education has become an important concern for OECD's activities, and PISA has become the key project in this context. While the OECD has only 34 countries, the total number of countries and economies that take part in PISA is now about 70.

Since the first PISA round, PISA has extended its ambitions, range and scope. While PISA is a test for 15-year olds, a "PISA for adults" (PIAAC) has also been launched, and a "PISA for Development" is also developed, targeting what is considered to be competencies needed for young people in developing countries. A "PISA for schools" test is also developed as a commercial product. This test enables schools and school districts to compare themselves on a "PISA-like" scale with each other and with nations that are "PISA-winners". Plans for measuring aspects of pre-school and kindergarten are also on the drawing board.

In short: The OECD now provides indicators, numbers and rankings from early childhood to pension age



► FIG. 1:
First words on PISA
leader Andreas
Schleicher's TED-
talk on PISA.

that countries can (and do!) use to shape their own policy from a global context.

This work is certainly of great value. But statistics does not only describe reality, it also shapes and defines reality. **The mean PISA-score of a country is used as the definition of the quality of the nation's school system, and PISA-point per dollar is used as an indicator for the efficiency of the school system.**

While the OECD does not have formal political power, it exerts influence through its reports, policy papers and expert advice. This influence is called "soft power", "governance by numbers" and "governance by comparison", and there is an abundance of literature providing details of how OECD exerts this power globally to define reality and exert its power (see e.g. Meyer & Benavot 2013).

In many countries new curricula have been introduced, caused by "PISA-shocks", (e.g. Norway, Denmark, Sweden, Germany and Japan). In many countries new national standards as well as new systems of obligatory national testing have been introduced. Some of these are directly influenced by PISA documents, as also proudly noted in a comprehensive report by the OECD itself (Breakspear 2012).

PISA in alliance with commercial interests

The strife for better test scores also serves commercial interests. Companies deliver products like tests and teaching materials that are supposed to increase scores, and cramming schools make substantial profit from preparing students to achieve higher test scores. The largest PISA contractor is the US-based non-profit assessment and measurement institution ETS. Maybe more important is that the world's largest commercial educational company, Pearson Inc. was involved in PISA 2015 and won the bid to develop the framework for PISA 2018. The joint press release from OECD and Pearson explains:

"Pearson, the world's leading learning company, today announces that it has won a competitive tender by the Organisation for Economic Co-operation and Development (OECD) to develop the Frameworks for PISA 2018. [...] The frameworks define what will be measured in PISA 2018, how this will be reported and which approach will be chosen for the development of tests and questionnaires." (OECD & Pearson 2014)

The partnership with PISA/OECD is also a strategic door-opener for Pearson "with 40,000 employees in more than 70 countries" into the global educational market. In company with the OECD, Pearson also produces "The Learning Curve", a ranking of nations according to a set of test-based indicators. PISA leader Andreas Schleicher sits on the Advisory Panel of The Learning Curve. These rankings get media coverage and further create anxiety among politicians and policymakers. The result is a further pressure toward doing "something" to climb the league tables.



◀ FIG. 2: "Schule macht dum". The first PISA-results created a crisis in Germany, and led to large investments in science and mathematics education. (Bad results may be good news for science and mathematics educators)

PISA scores: uncertainties and errors

When looking at the actual mean scores on PISA, one may notice that the majority of the OECD countries are in middle group, with rather small differences in means. If it was a bike race, one would have said most countries came in the "peleton", and got the same time. In fact, rather small changes in total score can shift the country's rankings 10 to 20 places.

Moreover, the uncertainty in the published PISA-score is substantial. Some of this is measurement sampling error. Additional uncertainties are related to the calculations of the published PISA scores that are based on the test results. The total PISA test consists of some 10 hours testing time, and each student gets about two hours test from this pool. The national PISA score and other population parameters are calculated by using a rather elaborate technique, using Rasch analysis and Item Response Theory (IRT). The assumptions about the properties of data that underlie the use of these techniques are questioned by many psychometrics experts. The method was actually strongly modified in the last round of PISA, causing dramatic changes in scores for some countries.

These uncertainties and possible sources for others errors are not well communicated in the PISA reports.

PISA, THE PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT

- Organized by the OECD (Organization for Economic Cooperation and Development)
- Testing every 3 years. Started in 2000, now 6 rounds
- Three main domains: reading, mathematics and science
- Testing national samples of 15 year olds in schools
- Data also from students' and principals' questionnaire
- Results from PISA 2015 published in December 2016
- 72 countries and economies take part
- Testing based on a framework ("curriculum") developed by appointed experts
- Test items not "school science"
- The test is digital/electronic from 2015

Experts on psychometry criticise PISA for downplaying the problematic nature of the calculations and the lack of openness regarding sources of error and methodological uncertainties. Experts also criticise PISA reports for drawing unwarranted conclusions, and urge OECD to have a "a more measured approach to reporting and interpreting PISA Results" (Rutkowski & Rutkowski 2016). Some scholars use even stronger words, arguing that PISA is a "Tale of flaws and hubris" (Meyer 2013).

PISA: Intriguing findings

While the PISA scores and country ranking receives a lot of attention, other results pass more or less unnoticed. Here are some of the intriguing problematic results:

- Money and resources spent on education do not seem to be positively related to PISA scores.
- Class size is not related to PISA scores.
- PISA-scores correlate negatively with investment in and the use of ICT in teaching.
- PISA science scores seem unrelated to the time given to science in school.
- There is clear negative correlation with a country's PISA test score and the construct "Interest in science". For instance, Finland was in PISA 2006 on top of the PISA science score and at the very bottom of the index for interest in science. (see figs below)
- "PISA-winners" (Japan, Korea, Taiwan, Shanghai, Finland) students report very little use of inquiry-based teaching, the kind of teaching is recommended by the EU (2007) as well as by ICSU (2011). IBSE (Inquiry-Based

Science Education) is also the key concept in Horizon 2020 to get funding for research and development in science education.

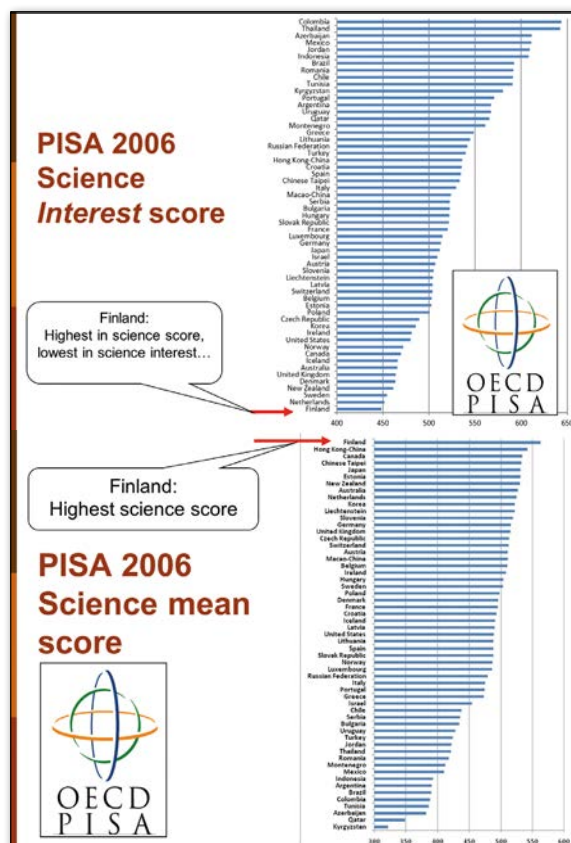
- Even for the variation within the same country, the PISA finding is that "in no education system do students who reported that they are frequently exposed to enquiry based instruction [...] score higher in science." (OECD 2016, p 36)
 - Experiments play a crucial role in science, and have always played an important role in science teaching at all levels. But when it comes to PISA, the recent report states that: "activities related to experiments and laboratory work show the strongest negative relationship with science performance" (OECD 2016, p 71)
- Whether one "believes in PISA" or not, such intriguing results need to be taken seriously and discussed. ■

About the Author



Svein Sjøberg is professor emeritus in physics education at Oslo University. Educated as a nuclear physicist, later also in education (MA) and educational psychology (PhD). Svein has been involved in expert committees on science education for national institutions and research councils in several countries as well for UNESCO, OECD, The European Commission and ICSU (International Council of Science). Has been awarded several national and international prizes and awards for his research and outreach activities..

► FIG. 3: Science was the main subject in PISA 2006. Finland came out on top of the PISA test score, but at the very bottom on the "Interest in science"-score.



References

- [1] S. Breakspear, (2012). *The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance*, OECD Education Working Papers, No. 71, OECD Publishing. <http://dx.doi.org/10.1787/5k9dfqf28-en>
- [2] EU (2007). *Science Education Now: A renewed pedagogy for the future of Europe*, (The Rocard report), European Commission, Brussels: EC.
- [3] H.D. Meyer and A. Benavot, (Eds) (2013). *PISA, Power and Policy: the emergence of global educational governance*, Oxford: Symposium Books
- [4] H.D. Meyer, (2013). *OECD's PISA: a Tale of Flaws and Hubris*. *TeachersCollege Record* (online), (www.tcrecord.org/content.asp?contentid=17371).
- [5] ICSU (2011). *Report of the ICSU Ad-hoc Review Panel on Science Education*. Paris: International Council for Science
- [6] OECD & Pearson (2014). *Pearson to develop PISA 2018 Student Assessment 21st Century Frameworks for OECD* (Joint press release) https://www.pearson.com/news/announcements/2014/december/pearson-to-develop-pisa-2018-student-assessment-21st-century-fra.html
- [7] OECD (2016). *PISA 2015 Results (Volume II): Policies and Practices for Successful Schools*. Paris: OECD Publishing.
- [8] L. Rutkowski & D. Rutkowski (2016). *A Call for a More Measured Approach to Reporting and Interpreting PISA Results*. *Educational Researcher*, Vol. 45 No. 4, pp. 252–257
- [9] A. Schleicher, (2013). *Use data to build better schools*. TEDGlobal http://www.ted.com/talks/andreas_schleicher_use_data_to_build_better_schools?language=en