

Notes for the talk of September 1st with Michele Catanzaro, European Conference of Science Journalism.

By Andrea Saltelli, August 31, 2020 Barcelona

Programme at: <http://www.ecsj2020.eu/programme/>

The Nature piece: [1]

The ecological footprint

This could be, for example, the case of the Environmental Footprint. As the time of writing the present article the 2019 edition of the measure – taken up by most media, just informed that in 2019 humans have overshoot 100 % of the planet yearly resources by July 29, and that by the end of the year humans activities shall have used 1.75 planets, instead of the one available. The critique that follows might appear to the reader one of those academic disputes which are all the most virulent the less important are the points of contention. In fact, several scholars of different orientation and discipline have noted that while it is evident that man is overexploiting natural resources, the metrification of this into numbers and dates (1.75 planets, July 29) as proposed by the Ecological Footprint does not withstand scrutiny. The details can be read in the exchange of papers written by both proponents and dissenters ([Galli et al., 2016](#); [Giampietro & Saltelli, 2014a, 2014b](#); [Goldfinger, Wackernagel, Galli, Lazarus, & Lin, 2014](#); [Blomqvist et al., 2013a, 2013b](#); [Rees & Wackernagel, 2013](#); [Van Den Bergh & Grazi, 2010](#)). Summarizing these critiques – which include the negative judgement of the international commission on the measurement of progress led by Nobel laureates Amartya Sen and Joseph Stiglitz with the French economist Jean-Paul Fitoussi ([Stiglitz, Sen, & Fitoussi, 2009](#)) would take much of the space available for the present work. Suffices to say that the 1.75 planets could easily be twenty or two hundred or infinity, if numbers have a meaning, depending on what impacts are measured and how; that the impact of human activity on the planet is too complex to be captured by a single number with three digits accuracy (1.75); and that the concept whereby the earth has a neat yearly budget which humans can use is incredibly optimistic against all those impacts which are irreversible – e.g. loss of species, depletion of non-renewable resources, persistent pollutants and so on. The Ecological Footprint is very effective in showing the urgency of action to reduce the pressure on the planet. Why should one be fastidious about the details? A number conveys an irresistible impression of accuracy and allows the setting of target. If humanity could take 5 days away every year from the overshoot date ([Global Footprint Network, 2019](#)) – as suggested by the proponents, things would go in the right direction. Unfortunately, as agreed by all critics – this measure is particularly ineffective for policy, e.g. in prioritizing what aspects of this pressure to reduce. If there must be an ethics of quantification which is not purely consequentialist, the seduction of the Ecological Footprint should be resisted. (Sorice: [2])

Shanghai ranking of world universities (ARWU)

The case of the ecological footprint is not unique. If one were to mention other extremely successful measures of recent times – by impact of society and academia alike, one would likely mention the Shanghai Rankings – known as Academic Ranking of World Universities

(ARWU). Neither the ecological footprint nor ARWU were developed by international organizations (such as the OECD, EC, WTO, WEF...), by a statistical office, by a pre-existing disciplinary stronghold or by a think tank. Their fortune owes to the work and the ingenuity of their developers. Though the flaws of ARWU are venial ([Paruolo, Saisana, & Saltelli, 2013](#); [Saisana, d’Hombres, & Saltelli, 2011](#)), before the audacious acrobatics of the ecological footprint just discussed, both measures have received pointed criticism from academia and practitioners. Note also that while the ecological footprint scores low on quality, it is not particularly damaging – one would assume, in its societal implications; it reminds humans inhabiting the developed world that they should consume less, and this admonition is couched into an overall non stressing – one might say optimistic, vision of humanity’s impact on the planet. To the extent that no actual policy is undertaken to match its scores, - e.g. that no country reduces its food import just to be seen consuming less ‘ghost land’ ([Giampietro & Saltelli, 2014a](#); [Stiglitz et al., 2009](#)), the Ecological Footprint measure makes media satisfied and the publics - in a sense, sedated. The ARWU, while apparently less ambitious in scope, has been rightly noted for its performativity: it forces the university system in the trajectories of a global market for education ([Éloire, 2010](#)). Thus, it has been labelled as societally damaging – a least by some observers.

As evident, the media appear to have little interest in technical disquisitions before the effectiveness of an appealing narrative. Thus, little of the existing criticisms has truly impacted the societal use and take up of these two metrics. This poses the question of who is responsible, when the public intellectuals populating media and academia apparently fail in their role. An interpretation of this conundrum in terms a media system overflowing the specific systems of science, technology and policy is offered in ([Saltelli & Boulanger, 2019](#)) using the theoretical implant of Luhmann’s social system theory. (Source: [2])

... en objectivant une hiérarchie qui, auparavant, était soit informelle soit inexistante, il contribue à générer artificiellement les conditions d’un marché des universités. Depuis le processus de Bologne (en 1999) et la stratégie de Lisbonne (en 2000) [[1](#)][Isabelle Bruno, À vos marques, prêts, cherchez ! La stratégie...](#), ce passage d’une logique de service public à une logique de marché, concurrentielle et gestionnaire, se présente comme inéluctable. Elle n’est, en fait, issue que d’une conception néolibérale du fonctionnement de la société, conception qui s’évertue à placer « l’économique » comme horizon unique et indépassable de l’humanité. La « politique de société » chère aux néolibéraux, consiste à « intervenir sur la société elle-même dans sa trame et dans son épaisseur [...] pour que les mécanismes concurrentiels, à chaque instant et en chaque point de l’épaisseur sociale, puissent jouer le rôle de régulateur [[2](#)][Michel Foucault, Naissance de la biopolitique. Cours au Collège...](#) »

... En imposant arbitrairement ses critères, le classement de Shanghai impose, hors de tout débat démocratique, une vision normative de ce qu’est une « bonne » université. Et, en ne laissant place à aucune tentative de proposer une alternative à cette vision d’une université « efficace » économiquement, le classement règle unilatéralement la question du rôle que doit jouer l’université dans la société. (Source: [3]).

See also [[4](#)][[5](#)].

Cost of climate action/inaction

Nicholas Stern argues that today's integrated assessment models for quantifying the economic and societal impacts of climate change are inadequate ([Nature 530, 407–409; 2016](#)). We disagree with his view on the superiority of more complex models such as DSGE (dynamic stochastic computable general equilibrium) models, which purport to account for a larger class of uncertain future events.

In our view, DSGE models have proved to be ineffective for policymaking, even in simple, short-term settings of pure economics, by failing to anticipate the onset of the recent recession (see P. Mirowski *Never Let a Serious Crisis Go to Waste* 275–286; Verso, 2013). Three decades of social-sciences research on science and politics make it clear that cost–benefit models cannot tame policy-relevant uncertainties or promote political agreement (see, for example, D. Collingridge and C. Reeve *Science Speaks to Power* 3–4, 59–60; Pinter, 1986).

Models that predict higher costs of climate change might make political intervention more palatable. But prescribing models that generate more precisely quantified estimates of a desired output is a political programme, not a scientific one. Responsible research requires responsible quantification and responsible acknowledgement of uncertainty. (Source: [6]).

OECD-PISA

[The Programme for International Student Assessment](#) (PISA) implemented by the OECD has been controversial since the publication of its first results in 2000.

Measuring the mathematics, science and reading skills of 15-year-old students every three years, PISA relies on broad international participation. In the 2015 test, as many as [72](#) countries joined the exercise, including those outside the OECD.

It's common to find articles where PISA is presented as a measure of a country's [innovation and growth potential](#). But it's also not rare to find others where the metrics used are [contested](#) as irrelevant and potentially counter-productive.

For advocates of “evidence-based” or “informed” policy, PISA incarnates the dispassionate, objective facts that nourish the formulation of good approaches – in this case, in the field of education. It allows for country comparisons and can help identify good practices that are worth emulating.

[Opponents](#) of the program reject the choice made by the OECD to link education and economic growth. While this choice was explained at the beginning as a way of showing [the high cost of a low educational performance](#), for some it also embodies [a neoliberal framing of education policies](#), which [forces the sector into the context of globalisation](#).

Post-normal science

In a new [review study](#) in the [International Journal of Comparative Education and Development](#) that I co-authored, we consider that facts such as those produced by PISA can

be viewed through the lense of [post-normal science](#). This approach is particularly apt for assessing scientific evidence when it feeds the policy process.

Post-normal science is a problem-solving strategy for issues where “facts are uncertain, values in dispute, stakes high and decisions urgent”. The concept was [created](#) in the 1990s by [Silvio Funtowicz](#) and Jerome R. Ravetz.

A key idea of the schema is “extended participation”, which suggests opening analyses to experts from different disciplines and forms of scholarship (one of the demands of [PISA’s critics](#)). It also points toward the active participation of relevant and legitimate stakeholders.

According to our review, a post-normal science reading of the PISA survey and its implications identifies a rich mix of methodological and ideological issues – in keeping with its tenet that the distinction between facts and values becomes problematic when the stakes are high.

PISA gives scores to participating countries so they can be ranked from best to worst for the skills measured, as well as measuring how they stand globally over all skills. Too much importance is being given to these scores and rankings, given the many non-transparent assumptions made by the OECD in their construction.

We don’t know, for example, how choices are made to include or exclude questions. There is also an issue about how many and which students participate in the test. The latter issue generates the so-called “non-response bias” and noticeably affects results.

Our review discusses the case of PISA non-response for England, where the bias turned out to be [twice the size](#) of the OECD declared standard error in 2003. This case illustrates how the results are much more uncertain and dependent on non-educational variables than it appears on a superficial reading.

In other words, the margin of error on the scores provided by the OECD is underestimated, and the ranking of countries from best to worst is more open to interpretation than one would understand from OECD analyses. To resolve this issue, the OECD should provide PISA users with a structured [sensitivity analysis](#) that takes all the variables in the ranking into account.

Ideally this analysis could be made by the users themselves, but this would only be possible if the OECD made all data available, which is not the case at present.

The worldview selected for the PISA analysis is also contentious. A main issue with PISA’s ambition to measure life skills needed to function in knowledge societies, for example, is that these skill are assumed to be the same across countries and cultures. Nor is it clear that all societies can safely be assumed to be destined to become “knowledge-oriented”.

Other fundamental questions emerged from our reading, too. Is it acceptable to see education as an input into growth? Does PISA “flatten” curricula – [narrowing our collective imagination regarding what education is and ought to be about](#) – and encourage focusing on a subset of educational topics at the expenses of others?

As noted by [our study](#), country comparison is achieved by “ignoring the great diversity of curricula across the participating countries – diversity which might in fact be a source of country-specific creativity and well-being”.

Urgency and caution

The PISA controversy is a helpful reminder that citizens in democracies must be critical of the facts that feed into public discourse. This [predates](#) any alarm about the emergence of the purported post-truth society, though that has made these discussions are [more urgent](#).

Facts must be taken with caution. For example, PISA scores have supported such inferences as [this paragraph in a 2014 study](#) prepared for the European Commission:

If every EU Member State achieved an improvement of 25 points in its PISA score (which is what for example Germany and Poland achieved over the last decade), the GDP of the whole EU would increase by between 4% and 6% by 2090; such an 6% increase would correspond to 35 trillion Euro.

The authoritative tone and use of crisp numbers here suggest causality – from education to growth – and an air of accuracy in a claim that is more like an act of faith than the result of scientific processes.

Our review of the PISA controversy also highlighted a problem of power in the use of evidence. With PISA, the OECD – an international organisation composed of unelected officers and scholars – has constructed a neoliberal framing of education policy and used its authority to dominate the global conversation, potentially at the expense of national or regional authorities and institutions.

This “global super-ministry of education”, in the words of an educator quoted in our study, effectively marginalises alternative visions of education that would normally hold weight. Thus the idea of education as personal development and fulfilment, what Germans call *Bildung*, becomes invisible, because it cannot be used as an internationally comparable metric.

Different skills

A full discussion of all points of controversy would take more space than this contribution allows, and should touch on the tension in using metrics to appreciate cognitive skills, as well as the need for other skills, such as [critical thinking, intrinsic motivation, resilience, self-management, resourcefulness, and relationship-building](#).

The OECD is unlikely to [suspend](#) or abolish the PISA study, in part because it serves a function. Before PISA, a country’s educational development was approximated by the average number of years of schooling there.

PISA raised awareness of other factors beyond classroom hours, such as literacy, that affect students’ educational outcomes. For those who study education, standardized tests like PISA also offer a useful instrument for comparing within and among countries.

[The Programme for International Student Assessment](#) (PISA) implemented by the OECD has been controversial since the publication of its first results in 2000.

Measuring the mathematics, science and reading skills of 15-year-old students every three years, PISA relies on broad international participation. In the 2015 test, as many as [72](#) countries joined the exercise, including those outside the OECD.

It's common to find articles where PISA is presented as a measure of a country's [innovation and growth potential](#). But it's also not rare to find others where the metrics used are [contested](#) as irrelevant and potentially counter-productive.

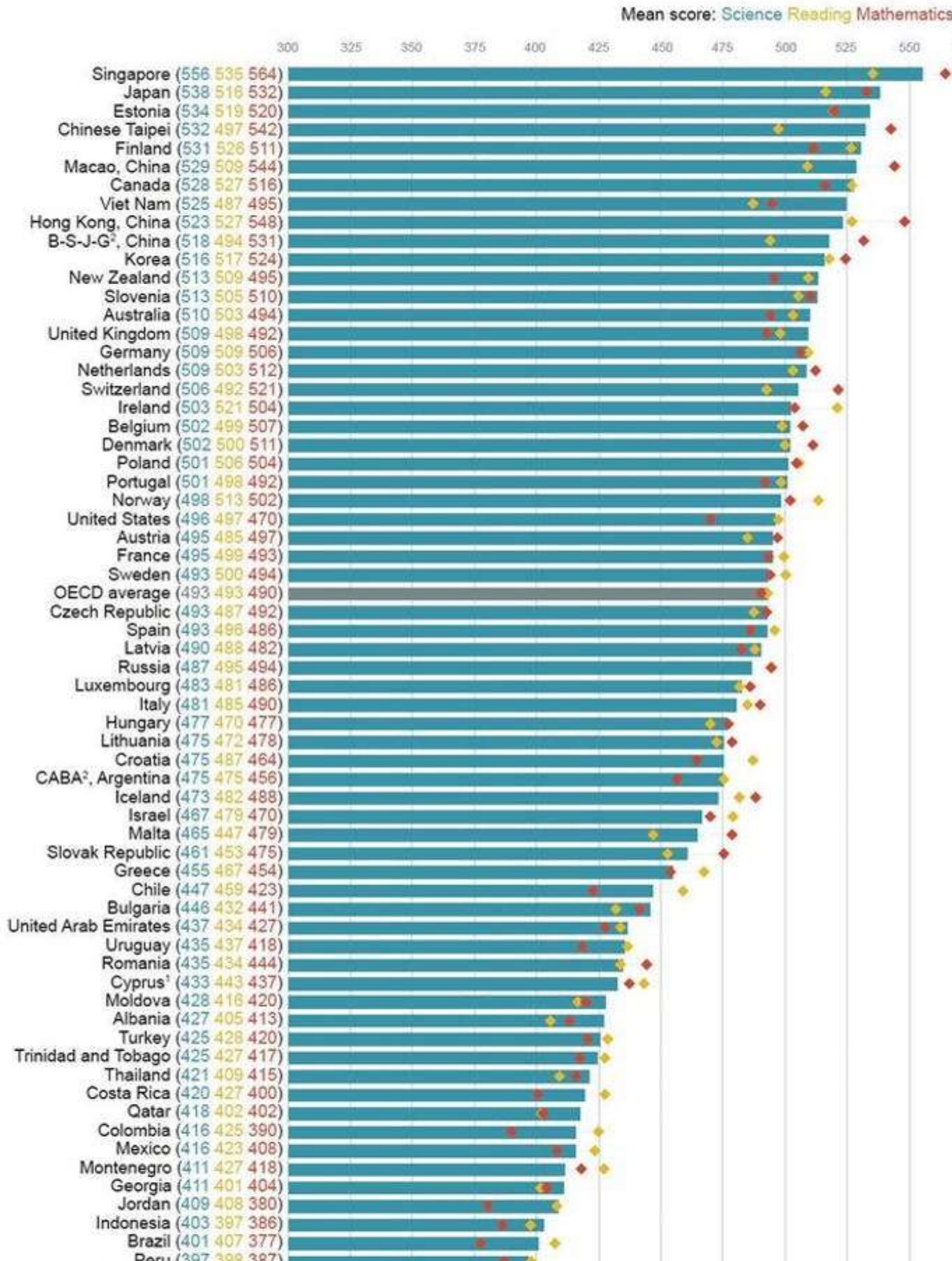
- [1] A. Saltelli *et al.*, "Five ways to ensure that models serve society: a manifesto," *Nature*, vol. 582, pp. 482–484, 2020.
- [2] A. Saltelli, "Ethics of quantification or quantification of ethics?," *Futures*, vol. 116, 2020.
- [3] F. Éloire, "Le classement de Shanghai. Histoire, analyse et critique," *L'Homme la société*, vol. 178, no. 4, p. 17, 2010.
- [4] C. O'Neil, *Weapons of math destruction : how big data increases inequality and threatens democracy*. Random House Publishing Group, 2016.
- [5] J. Z. Muller, *The tyranny of metrics*. Princeton University Press , 2018.
- [6] A. Saltelli, S. Funtowicz, M. Giampietro, D. Sarewitz, P. B. Stark, and J. P. van der Sluijs, "Modelling: Climate costing is politics not science," *Nature*, vol. 532, no. 7598, pp. 177–177, Apr. 2016.
- [7] A. Saltelli, "International PISA tests show how evidence-based policy can go wrong," *The Conversation*, Jun-2017.



PISA 2015 results



Snapshot of performance in science, reading and mathematics



PISA results for 2015. [OECD](#)

For advocates of “evidence-based” or “informed” policy, PISA incarnates the dispassionate, objective facts that nourish the formulation of good approaches – in this case, in the field of education. It allows for country comparisons and can help identify good practices that are worth emulating.

[Opponents](#) of the program reject the choice made by the OECD to link education and economic growth. While this choice was explained at the beginning as a way of showing [the high cost of a low educational performance](#), for some it also embodies [a neoliberal framing of education policies](#), which [forces the sector into the context of globalisation](#).

Post-normal science

In a new [review study](#) in the [International Journal of Comparative Education and Development](#) that I co-authored, we consider that facts such as those produced by PISA can be viewed through the lense of [post-normal science](#). This approach is particularly apt for assessing scientific evidence when it feeds the policy process.

Post-normal science is a problem-solving strategy for issues where “facts are uncertain, values in dispute, stakes high and decisions urgent”. The concept was [created](#) in the 1990s by [Silvio Funtowicz](#) and Jerome R. Ravetz.

A key idea of the schema is “extended participation”, which suggests opening analyses to experts from different disciplines and forms of scholarship (one of the demands of [PISA’s critics](#)). It also points toward the active participation of relevant and legitimate stakeholders.

According to our review, a post-normal science reading of the PISA survey and its implications identifies a rich mix of methodological and ideological issues – in keeping with its tenet that the distinction between facts and values becomes problematic when the stakes are high.

PISA gives scores to participating countries so they can be ranked from best to worst for the skills measured, as well as measuring how they stand globally over all skills. Too much importance is being given to these scores and rankings, given the many non-transparent assumptions made by the OECD in their construction.

We don’t know, for example, how choices are made to include or exclude questions. There is also an issue about how many and which students participate in the test. The latter issue generates the so-called “non-response bias” and noticeably affects results.

Our review discusses the case of PISA non-response for England, where the bias turned out to be [twice the size](#) of the OECD declared standard error in 2003. This case illustrates how the results are much more uncertain and dependent on non-educational variables than it appears on a superficial reading.

In other words, the margin of error on the scores provided by the OECD is underestimated, and the ranking of countries from best to worst is more open to interpretation than one would understand from OECD analyses. To resolve this issue, the OECD should provide PISA users with a structured [sensitivity analysis](#) that takes all the variables in the ranking into account.

Ideally this analysis could be made by the users themselves, but this would only be possible if the OECD made all data available, which is not the case at present.

The worldview selected for the PISA analysis is also contentious. A main issue with PISA's ambition to measure life skills needed to function in knowledge societies, for example, is that these skills are assumed to be the same across countries and cultures. Nor is it clear that all societies can safely be assumed to be destined to become "knowledge-oriented".

Other fundamental questions emerged from our reading, too. Is it acceptable to see education as an input into growth? Does PISA "flatten" curricula – [narrowing our collective imagination regarding what education is and ought to be about](#) – and encourage focusing on a subset of educational topics at the expenses of others?

As noted by [our study](#), country comparison is achieved by "ignoring the great diversity of curricula across the participating countries – diversity which might in fact be a source of country-specific creativity and well-being".

Urgency and caution

The PISA controversy is a helpful reminder that citizens in democracies must be critical of the facts that feed into public discourse. This [predates](#) any alarm about the emergence of the purported post-truth society, though that has made these discussions are [more urgent](#).

Facts must be taken with caution. For example, PISA scores have supported such inferences as [this paragraph in a 2014 study](#) prepared for the European Commission:

If every EU Member State achieved an improvement of 25 points in its PISA score (which is what for example Germany and Poland achieved over the last decade), the GDP of the whole EU would increase by between 4% and 6% by 2090; such an 6% increase would correspond to 35 trillion Euro.

The authoritative tone and use of crisp numbers here suggest causality – from education to growth – and an air of accuracy in a claim that is more like an act of faith than the result of scientific processes.

Our review of the PISA controversy also highlighted a problem of power in the use of evidence. With PISA, the OECD – an international organisation composed of unelected officers and scholars – has constructed a neoliberal framing of education policy and used its authority to dominate the global conversation, potentially at the expense of national or regional authorities and institutions.

This "global super-ministry of education", in the words of an educator quoted in our study, effectively marginalises alternative visions of education that would normally hold weight. Thus the idea of education as personal development and fulfilment, what Germans call *Bildung*, becomes invisible, because it cannot be used as an internationally comparable metric.

Different skills

A full discussion of all points of controversy would take more space than this contribution allows, and should touch on the tension in using metrics to appreciate cognitive skills, as well as the need for other skills, such as [critical thinking, intrinsic motivation, resilience, self-management, resourcefulness, and relationship-building](#).

The OECD is unlikely to [suspend](#) or abolish the PISA study, in part because it serves a function. Before PISA, a country's educational development was approximated by the average number of years of schooling there.

PISA raised awareness of other factors beyond classroom hours, such as literacy, that affect students' educational outcomes. For those who study education, standardized tests like PISA also offer a useful instrument for comparing within and among countries.

Still, our study reinforces that democratic societies view "evidence-based policy" with a critical eye, querying who produced the evidence and whose interests are served by it. PISA is a strong example of the [power asymmetries](#) inherent in producing facts to inform policy.

In this case the OECD, possibly the most muscular player in the arena of international education policy, can frame evidence around its preferred norms and impose them on public discourse. (Source: [7]; see also https://www.elconfidencial.com/alma-corazon-vida/2019-12-03/no-deberiamos-dar-importancia-pisa-saltelli_2358835/).

- [1] A. Saltelli *et al.*, "Five ways to ensure that models serve society: a manifesto," *Nature*, vol. 582, pp. 482–484, 2020.
- [2] A. Saltelli, "Ethics of quantification or quantification of ethics?," *Futures*, vol. 116, 2020.
- [3] F. Éloire, "Le classement de Shanghai. Histoire, analyse et critique," *L'Homme la société*, vol. 178, no. 4, p. 17, 2010.
- [4] C. O'Neil, *Weapons of math destruction : how big data increases inequality and threatens democracy*. Random House Publishing Group, 2016.
- [5] J. Z. Muller, *The tyranny of metrics*. Princeton University Press , 2018.
- [6] A. Saltelli, S. Funtowicz, M. Giampietro, D. Sarewitz, P. B. Stark, and J. P. van der Sluijs, "Modelling: Climate costing is politics not science," *Nature*, vol. 532, no. 7598, pp. 177–177, Apr. 2016.
- [7] A. Saltelli, "International PISA tests show how evidence-based policy can go wrong," *The Conversation*, Jun-2017.