

Thinking critically about PISA

ROSEMARY HIPKINS

PISA will be in the news again this year. The 2018 results are due to be released at the end of 2019 and they usually generate media interest. This Rangahau Whakarāpopoto is a research brief which outlines things to watch out for as you think about what the results might mean.

What is PISA?

PISA is an acronym for the Programme for International Student Assessment. It is owned by the Organization for Economic Co-operation and Development (OECD) and is the largest international study of education ever undertaken. PISA's purpose is to assess the readiness of 15 year olds to participate in life and work when they leave school. This purpose differs from the more traditional assessment focus on knowing things. Instead, PISA aims both to assess what students can do with what they know, and to relate these results to the education system of the country in which they have been educated. The OECD developed the original version of the key competencies which were designed to help shape assessments that could help achieve these stated aims for PISA. The key competencies in *The New Zealand Curriculum* (Ministry of Education, 2007) (NZC) were informed by this work. The OECD is now working on an updated version called the 2030 model (OECD, 2018a). This new model will likely impact the design of future PISA assessments.

Why does PISA matter?

PISA has made OECD a very powerful player in the education space. PISA has become an international

yardstick of the success of education systems, and governments don't want to be left out. Increasingly, nations that don't belong to the OECD are also opting in. Nations pay to participate in PISA and fund the data collection in their own countries.

PISA assessments generate many reports that create debates about education systems:

- Some reports are international comparisons. PISA 2015 Results in Focus is an example from the 2015 round (OECD, 2018b). By the time you are reading this article, early 2018 reports should be available.
- Some reports are specific to one nation: these tend to be written by officials in that nation, drawing on the large database of results. The example I have chosen is a report specific to New Zealand's performance in the 2015 PISA round (Ministry of Education, 2016).
- Some reports discuss design challenges related to the sorts of thing PISA endeavours to assess, given the agenda just outlined. The example I have chosen explains how the OECD planned to go about assessing "global competence" in 2018 (OECD, 2018c).
- Publications in the OECD's PISA in Focus series are "quick reads" for practitioners. Each addresses a specific small slice of PISA-related information.

What sort of research information does PISA give us?

PISA began with the intention to assess English literacy, mathematics literacy, and science literacy. In each 3-year cycle one of these will be a major focus and the other two will be assessed less thoroughly. The emphasis on “literacy” in science and maths signals the intention to assess students’ ability to apply learning to a range of life contexts. In practice, this is easier said than done but it is a point of difference from TIMSS (The International Mathematics and Science Study), which has a more-traditional assessment focus on content and thinking skills.

Over the years since PISA began, the OECD has added new assessment focuses. One example is collaborative problem-solving (OECD, 2017). Another is global competencies. These assessments include social-emotional elements that are not easy to assess. The OECD prepares discussion documents—like the one about global competencies—to attempt to scope what experts call the “assessment construct”. This construct is what they will use as the basis for developing assessment items. The validity of these new assessment constructs is often controversial. Experts doubt they can assess what they say they want to assess, given the limitations of the assessment programme per se (more about this in a minute).

OECD has also been actively innovating in the area of e-assessment. Using the investment funds which the programme attracts from so many governments, they can afford to buy international expertise for the things they want to try out. Assessing students’ ability to collaborate virtually is one innovation they have tried that has been subject to critique by assessment experts (Shaw & Child, 2017).

As part of the assessment programme, PISA also gathers data both from students as to how they feel about their learning, and from teachers and school leaders about the opportunities they provide to students. The demographic data that are gathered include students’ nationalities and a measure of the relative economic wellbeing of their family. All these other data can be correlated with the achievement data in many different ways, and national reports will often dig deeper into relationships between achievement and the other variables. I have chosen an example that digs into contextual differences related to students’ science achievement in 2015 (Kirkham, with May, 2016).

International reports rank nations on a range of criteria. These are the reports that tend to become the focus of sensationalist commentary in the news media.

Why are there concerns about PISA?

Assessment researchers have expressed a range of concerns about PISA. Some concerns relate to how the results are used. For example, even minor negative changes in international rankings are often sensationalised by news media as a failing of the national education system, even though the actual reports are careful to group clusters of nations that cannot be meaningfully statistically separated in their performance. Too much is made of small differences, just as happens with data misuse in many different contexts.

Other concerns about the PISA assessment programme include the following.

- Even though there is a strict protocol to follow, studies have shown that translating the same question into different languages can change what is actually being asked in subtle ways that make the question comparatively easier or harder to answer. For example, Sjöberg (2015) explores the translation challenge in the context of Scandinavian languages.
- The need to use contexts that will make sense to students in many different nations severely constrains the actual contexts that can be used in the assessment items. It is also likely that some contexts will still give a comparative advantage to students in some nations. The OECD itself recognises this problem, and uses an analysis process called differential item functioning to check for such differences (Cresswell, Schwantner, & Waters, 2015).
- As already noted, the validity of some assessments has been challenged by experts—do they actually assess what they say they assess?
- Some critics claim that PISA helps generate the very inequalities it then reports. The concept of “consequential validity” is important here—basically the consequences of an assessment should be part of any analysis of its overall validity. South American researchers Taut and Palacios (2016) provide a detailed technical critique of the way that consequential validity is established for PISA assessments. They say that a “culture of blame” can develop when teachers and school leaders are unfairly blamed for disappointing PISA results, and that there is potential for PISA data to be misused to “label” groups of students who have not done well.
- PISA statisticians use psychometric modelling processes to fill in gaps in the data and to try to ensure comparability across different systems. There is a lot of technical critique of those modelling processes. A New Zealand Listener article (Woulfe, 2013) explores these issues from a New Zealand perspective. The article delves into multiple challenges for making fair comparisons, and outlines indications that the mathematical modelling used for this purpose is problematic.

- OECD gives advice about what nations should do next on the basis of their PISA outcomes. Some commentators are concerned that there is a drive to make everyone's education systems basically the same, regardless of differences in national values. Sjøberg (2015) makes this point, as does an article published on *The Conversation* blog site (Saltelli, 2017). Other commentators express concerns about the potential for perverse incentives—will constructive policy targets be created, or will policy makers look for short-term rankings gains (Breakspear, 2014)?

A letter to the PISA director from academics all around the world (including some from New Zealand) covers most of the issues listed above and adds one more—the potential for PISA to harm students' wellbeing, increasing stress levels for teachers and students alike (The Guardian, 2014).

Given these concerns, what should we take notice of?

Given the many different types of critique, we shouldn't get too panicked about minor changes in our PISA rankings. But PISA does have some useful things to tell us about our own education system. Researchers who write about PISA in a balanced way (see the reference below to the book *The Global Education Race*) say we should focus on data that throw up valid challenges within our own education system. For New Zealand, these challenges include:

- PISA data consistently show New Zealand as having one of the widest spreads of achievement of all the nations in the programme. Our best students are right up with the best internationally, but the same is also true of those who are underachieving (the phrase “the long tail of underachievement” has been used to describe this).
- The wide spread of achievement also applies within individual schools. We tend to rank high on PISA measures of inequality in our system overall.
- Demographic analyses show that Māori and Pasifika students are more likely to underachieve, compared with Pākehā or Asian students. We don't need PISA to tell us that of course—we can see similar patterns in the National Certificates of Educational Achievement (NCEA) and in NMSSA (the National Monitoring Study of Student Achievement).
- Over recent rounds, compared with the past performance of our students, changes in overall achievement levels have been negative. Again, we need to be careful about small changes (setting a comparable standard of difficulty when tests come 3 years apart is almost impossible). But a clear trend of falling achievement is of concern.
- PISA results can be analysed for indications of specific curriculum strengths and weaknesses of New Zealand students, relative to students in other nations.

For example, this summary report shows that statistics is an area of comparative strength for our students, but geometry and algebra are areas of comparative weakness (Ministry of Education, n.d.).

One report from the 2015 PISA round focused on student wellbeing (Ministry of Education, 2017). The report noted that higher levels of bullying were reported by New Zealand students compared with other nations. Findings like this can help inform initiatives to keep all students safe and enhance their wellbeing. As one specific example, the PISA report (Ministry of Education, 2017) has been linked to the “Pink Shirt” initiative which aims to address bullying. This example illustrates how the same factor—in this case wellbeing—can be seen as a concern (PISA can have negative impacts on wellbeing) and as support for action (we need to reduce bullying). The tricky thing is to use the findings to improve education for all our young people while minimising the potential for harm.

Want to follow up?

A brief such as this can only scratch the surface of complex issues. If you want to read more, this book would be a good place to start:

- Rutkowski, D., Thompson, G., & Sellar, S. (2017). *The global education race: Taking the measure of PISA and international testing*. Calgary, Canada: Brush Education. This short, easy-to-read book provides an excellent summary of the issues. If you would prefer to listen rather than read, there is a half-hour podcast about the book here: <https://soundcloud.com/freshed-podcast/freshed-74-the-global>

If you are interested in a systematic analysis of the huge body of PISA-related research, this article is a great place to start:

- Hopfenbeck, T., Lenkeit, J., El Masri, J., Cantrell, K., Ryan, J., & Baird, J. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the Programme for International Student Assessment. *Scandinavian Journal of Educational Research*, 62(3), 333–353. <https://doi.org/10.1080/00313831.2016.1258726>

References

- Breakspear, S. (2014). *How does PISA shape education policy-making? Why how we measure learning determines what counts in education*. Seminar Series Paper No. 240. Melbourne, VIC, Australia: Centre for Strategic Education. <https://simonbreakspear.com/wp-content/uploads/2015/09/Breakspear-PISA-Paper.pdf>
- Cresswell, J., Schwantner, U., & Waters, C. (2015). *A review of international large-scale assessments in education: Assessing component skills and collecting contextual data*. Paris: OECD.

- https://read.oecd-ilibrary.org/education/a-review-of-international-large-scale-assessments_9789264248373-en#page1
- Kirkham, S., with May, S. (2016). *PISA 2015: The science context for PISA*. Wellington: Ministry of Education. https://www.educationcounts.govt.nz/_data/assets/pdf_file/0017/180611/PISA-2015-Science-Context_v2.pdf
- Ministry of Education. (2016). *PISA 2015: New Zealand summary report*. Wellington: Author. <https://www.educationcounts.govt.nz/publications/series/PISA/pisa-2015/pisa-2015-summary-report>
- Ministry of Education. (2017). *PISA 2015: New Zealand students' wellbeing report*. Wellington: Author. https://www.educationcounts.govt.nz/_data/assets/pdf_file/0005/181544/PISA-2015-NZ-Students-Wellbeing-Report.pdf
- Ministry of Education. (n.d.). *Mathematics achievement: What we know from New Zealand's participation in TIMSS 2014/15 and PISA 2015*. Wellington: Author. Retrieved from https://www.educationcounts.govt.nz/_data/assets/pdf_file/0003/181569/Mathematics-achievement-What-we-know-from-NZs-participation-in-TIMSS-2014-15-and-PISA-2015.pdf
- OECD. (2017). *PISA 2015 results (Volume V) Collaborative problem solving*. Paris: Author. Retrieved from <http://www.oecd.org/education/pisa-2015-results-volume-v-9789264285521-en.htm>
- OECD. (2018a). *The future of education and skills: Education 2030*. Paris: Author. Retrieved from [https://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf)
- OECD. (2018b). *PISA 2015: PISA results in focus*. Paris: Author. Retrieved from <https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf>
- OECD. (2018c). *PISA preparing our youth for an inclusive and sustainable world: THE OECD PISA global competence framework*. Paris: Author. Retrieved from <https://www.oecd.org/pisa/Handbook-PISA-2018-Global-Competence.pdf>
- Saltelli, A. (2017, 12 June). International PISA tests show how evidence-based policy can go wrong. *The Conversation*. Retrieved from <https://theconversation.com/international-pisa-tests-show-how-evidence-based-policy-can-go-wrong-77847>
- Shaw, S., & Child, S. (2017). *Utilising technology in the assessment of collaboration: A critique of PISA's collaborative problem-solving tasks*. Cambridge, UK: Cambridge Assessment. Retrieved from <https://www.cambridgeassessment.org.uk/Images/424748-utilising-technology-in-the-assessment-of-collaboration-a-critique-of-pisa-s-collaborative-problem-solving-tasks.pdf>
- Sjøberg, S. (2015). PISA and global educational governance—A critique of the project, its uses and implications. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(1), 111–127. <https://doi.org/10.12973/eurasia.2015.1310a>
- Taut, S., & Palacios, D. (2016). Intended and unintended interpretations and uses of PISA results: A consequential validity perspective. *Revista Electronica de Investigacion y Evaluacion Educativa*, 22(1), 1–15. Retrieved from https://www.uv.es/RELIEVE/v22n1/RELIEVEv22n1_M8eng.pdf
- The Guardian. (2014, 6 May). OECD and Pisa tests are damaging education worldwide—academics. *The Guardian*. Retrieved from <https://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics>
- Woulfe, C. (2013, 4 December). Education rankings “flawed”. *New Zealand Listener*. <https://www.noted.co.nz/archive/listener-nz-2013/education-rankings-flawed/>

Dr Rosemary Hipkins is a chief researcher at NZCER.

Email: Rose.Hipkins@nzcer.org.nz