From your experience as a former Director at OECD, how has PISA changed how countries approach educational reform and system-level change?

A major feature of the work of the Organisation for Economic Co-operation and Development (OECD) is the collection of statistics and the development indicators as the basis for policy development and debate. Prior to 2000, its education statistics and indicators predominantly focused on inputs. The outcomes included were completion rates of various levels of schooling, graduation rates from tertiary education, and some labor market data on participation, employment, and unemployment rates for various age groups.

The proxy for quality of education, used as the measure of ‘human capital,’ was the number of years of formal education completed. Its use required the rather heroic assumption that a given number of years of formal education produced the same quality of ‘human capital’ in all countries. The one exception was direct data on the quality of students’ learning obtained for those OECD countries that participated in surveys conducted by the International Association for the Evaluation of Educational Achievement (IEA).

In the 1990s, the member countries of the OECD decided to gather their own data on students’ learning in what became the Program for International Student Assessment (PISA). Fewer than half the OECD member countries had committed when the program was established but, when the program was implemented for the first time in 2000, all but one of the 29 member countries and four others participated. A further 11 collected data in 2001 using the PISA 2000 tests. In the fourth PISA survey in 2009, there were 65 participants with a further 10 collecting data in 2010.

Reading, mathematics, and science were chosen as the domains to be tested but, unlike the IEA surveys, the assessment frameworks were not developed from an analysis of the curricula in the participating countries but rather from a specification of what it was students might be expected to be able to do with what they have learned. Some
countries expressed concern about the validity of such tests on the grounds that their curricula did not focus on developing such skills. In the end, all agreed that, however their curricula were formulated, there was a common intention that students should be able to use what they have learned and not just be able to demonstrate that they have learned it and that this was an appropriate criterion for judging quality.

The population chosen to be surveyed were 15-year-olds, the highest age at which virtually the whole cohort was still in formal education in OECD countries. Samples were age-based rather than grade-based since school starting ages vary. A similar choice was made when the U.S. National Assessment of Educational Progress (NAEP) was established for the same reason but a parallel, more administratively convenient grade-based sample was added in the 1980s and later, the age-based sample was dropped.

With PISA’s direct measures of the quality of learning, the OECD has fundamentally changed much of the discussion in international comparative education and in national education policy. With these measures of quality, a much stronger relationship between human capital and economic development has been established (e.g. Hanushek & Woessmann, 2008).

PISA has also provided a richer basis for more general policy studies. Disappointed with it’s initial PISA results led Denmark to commission an OECD (2004) review that included a detailed comparison with Canada (Ontario), Finland, and the United Kingdom.

The profile of PISA has also stimulated debate about its epistemology and methodology and whether it is imposing a uniformity shaped by a narrow view of the outcomes (e.g. Rizvi & Lingard, 2010).

In educational policy discussions in Australia, it has frequently been asserted that a pursuit of equity must be at the expense of quality. When I conducted a review of the New South Wales Higher School Certificate, the Grade 11-12 curriculum and its associated public examination system on which school graduation and university entrance are based, I encountered this view.

The existing curriculum was highly differentiated with the ostensible aim of catering appropriately for student diversity. The standard model was for three hierarchically arranged courses in each subject though there were four in English and five in mathematics. An analysis of Grade 11-12 enrollments in government schools in two major regions of Sydney showed that, in the more disadvantaged region, students who had been in the top 10 percent in English in Grade 10 were more likely to be enrolled in one of the two least demanding English courses in Grade 11-12 whereas students in the more advantaged region were more likely to be enrolled in one of the two most demanding courses. The differences in enrollment patterns reflected opportunities. Public schools in the more disadvantaged region were less likely to offer the more demanding courses (McGaw, 1997, pp. 445-45). What was a highly differentiated set of offerings at the state-level was not a reality in schools. Many strong students in areas perceived to be disadvantaged were denied access to demanding courses.

When I proposed a reduction in the extent of course differentiation this was attacked as a prescription for ‘dumbing down’ the curriculum. I countered that with a proposal for ‘leveling up’ and, in the end, was successful in persuading the government.
to take this approach. In the subsequent decade, increased differentiation has been gradually reintroduced through the addition of ‘extension courses.’

The first PISA report showed that there were countries that produced high quality and high equity (OECD, 2001, p. 253). The PISA tests provided the quality measures tests. The equity measure was the slope of the regression line for within-country analyses of the relationship between students’ social backgrounds and their performances on PISA. For all countries there was a positive slope, or social gradient, indicating that more advantaged students generally performed better than less advantaged ones but the slopes varied markedly across countries. Some of the high-performing countries, such as Canada, Finland, Japan, and Korea had social gradients significantly less steep than that for the OECD as a whole. Other high-performing countries such as Australia had social gradients significantly steeper than that for the OECD as a whole. They could be characterized as high-quality, low-equity countries. These results helped shift the debate in Australia from a ‘trade-off’ between quality and equity to the pursuit of both together.

The PISA surveys have consistently shown that around 70 percent of the difference between Australian schools in performance can be accounted for in terms of the differences in the social background of the students they enroll (OECD, UNESCO, 2003, p. 357). There is no formal policy in Australia to differentiate schools, as there is in Germany, for example. In Australia, the differentiation arises from demographic differences between geographic regions and the availability of a non-governmental school sector that enrolls one third of all students.

That structure of public and private provision of schooling in Australia is well established. The policy question is not how the structure might be changed but how social inequities might be reduced within it.

You chair the Australian Curriculum, Assessment and Reporting Authority (ACARA). How is ACARA connecting the national assessment program with the national curriculum?

The Australian National Assessment Program has two components. One currently involves a three-year cycle of sample-based surveys of students’ achievements in science literacy, civics and citizenship, and ICT literacy. Other sample-based surveys could be introduced as part of a program to monitor the impact of the new curricula now beginning to be rolled out.

The other is the National Assessment Program: Literacy and Numeracy (NAPLAN), which involves full-cohort assessments of students in Grades 3, 5, 7, and 9 in literacy and numeracy. This program arose from separate state and territory assessments of literacy and numeracy, the first of which was implemented in New South Wales as its Basic Skills Testing Program in 1990. For a period, the results of the separate jurisdictional assessments were mapped onto common scales but the state, territory, and federal ministers for education resolved to adopt common assessments across the country. This began as NAPLAN in 2008.

The NAPLAN assessments were essentially an amalgam of the different assessment practices of the states and territories. The Australian Curriculum will provide the basis for aligning the NAPLAN assessments with the curriculum as it is fully implemented. The current intention is to produce redeveloped NAPLAN assessments for use in 2014.
We are examining the possibility of electronic delivery of NAPLAN from 2014. This would allow the tests to be tailored to provide more precise estimates of individual students. It would also reduce ceiling and floor effects in the current tests.

We are also considering ways in which the tests might be broadened. In sample-based assessments such as PISA, the assessment domain is fully defined and assessment materials that cover it are developed before consideration is given to how much of the material a student could deal with in the time allocated for assessment. With a randomized block design, all assessment materials are used and the full domain assessed with individual students taking only part of the assessment. We are considering taking that approach with the NAPLAN tests used for full cohorts.

That would enable us to better represent the full scope of literacy and numeracy in the assessments. It would reduce the risk of the tests narrowing the curriculum both because the total scope of the assessments would be broader and because individual teachers would know that their students would be taking a variety of different components of the overall assessment materials. Any ‘teaching to the test’ would require teaching to develop well-rounded literacy and numeracy.

In what ways is the Measurement Framework for Schooling in Australia used to improve practice on the school level?

The NAP sample assessments and the results from PISA and IEA international assessments do not provide systematic information at the school level because of the nature of their samples. They inform system-level policy debate and development.

NAPLAN does provide information at the school level but there has been contention about how results at the school level might be interpreted. There is always a temptation to claim that relatively poor results are due to the characteristics of the students and so, in a sense, to claim that demography determines destiny.

To overcome this tendency schools are offered comparisons of their NAPLAN results with those of other schools with similar students. Using information on parents’ education and occupation, an index of socio-educational advantage has been developed. All schools are located on this scale and comparisons for each school are provided with 60 other schools, the 30 immediately above and the 30 immediately below on the scale.

These comparisons reveal remarkably different results among schools working in similar circumstances. They are essentially comparisons of ‘value added’ but without the highly contrived statistical estimates that can be difficult for lay users and even professional teachers to interpret.

The comparisons are made public on the My School website, so they are not restricted to the professional staff of a school. The information can inform parents’ choices of schools but, more importantly, it can inform parents better than ever before about how their school is performing. It can open a conversation between a school and its community about the school’s development program. It can open a conversation between schools about the policies and practices followed in the higher performing schools in a school’s comparison group of 60. Since these 60 can be from anywhere in the country, the conversations can be across geographic and jurisdictional boundaries and not necessarily between schools that might see themselves as local competitors.

This kind of use of school data has been undertaken for government schools within states and territories for some time. What is
new is that the comparisons are among all schools and that the results are publicly disclosed.

What do you see as one of or the most pressing issue related to educational change today?

For school education in Australia, I see the most pressing needs to be to improve on its already high quality while also reducing the impact of differences in students’ social background on their educational outcomes.

Educational change is a complex process and requires an array of strategies. Fundamental skills in literacy and numeracy, on which further learning depends, must be securely developed for all students, but a much broader base must also be built as the foundation for the development of individual expertise.

A curriculum with high expectations is one component. Systematic monitoring of performance is another. The highest priority, perhaps, should go to the preparation, recruitment, and retention of high quality teachers.

References


