Question Mark Over the Accuracy and Reliability of PISA Tests

The OECD's Programme for International Student Assessment (PISA) has extraordinary status and influence. It is seen as the gold standard for assessing the performance of education systems, but it is a castle built on sand. New data published by the Australian Council for Educational Research (ACER) call into question the accuracy and reliability of PISA and its league tables of country results.

The new figures show that nearly three-quarters of Australian students didn't fully try on the PISA 2018 tests. The ACER research found that "...the majority of Australian students (73%) indicated that they would have invested more effort if the PISA test counted towards their marks".

The ACER research also found that only 56% of students who participated in PISA 2018 put in high effort, 40% said they only put in medium effort and 5% said they put in low effort (figures don't add to 100% due to rounding). However, 91% said they would put in high effort if the tests counted towards their school marks.

These are remarkable revelations. How is it possible to accept the PISA results as an accurate measure of Australia's education performance if three-quarters of students didn't fully try in the tests?

There was also large variation between countries in the proportion of students that did not fully try on the tests. For example, 80% of German students and 79% of Danish and Canadian students didn't fully try on the tests compared to 38% of students in the four Chinese provinces participating in PISA, 45% in Taiwan and 46% in Korea. Across the OECD countries, the average proportion of students that made less than full effort was 68%.

The ACER also found significant differences on this measure between demographic groups in Australia. For example, 65% of Indigenous students didn't fully try compared to 74% of non-Indigenous students; 70% of provincial and remote area students didn't fully try compared to 75% of metropolitan students and 77% of females didn't fully try compared to 70% of males.

The most interesting difference was that 78% of socio-economically advantaged students did not fully try compared to 66% of disadvantaged students. This may be a factor in the decline in results amongst advantaged students in the tests over the past decade or so.

These results suggest that PISA is not the accurate, reliable, and valid measure of educational quality it claims. As the renowned international education scholar, Yong Zhao, observes, the PISA tests are an international education juggernaut that created "false idols of educational excellence for the world to worship".

The variation between countries in the proportion of students not fully trying also calls into question the validity of league tables of countries based on PISA results which attract so much publicity and commentary. Rankings can move up and down depending on student effort. The OECD has acknowledged that differences in student effort across countries will affect country results and rankings [p. 198].

A recent study published by the US National Bureau of Economic Research based on data from PISA 2015 found that differences in the proportion of students not fully trying had a large impact on the rankings for several countries. For example, it estimated that Portugal's ranking in science would have improved by 15 places from 31st to 16th if students had fully tried. Sweden's ranking would have improved 11 places from 33rd to 22nd and Australia's ranking by four places from 16th to 12th. It concluded:

Using PISA scores and rankings as done currently paints a distorted picture of where countries stand in both absolute and relative terms

These results raise the issue of whether changes in student effort over time contributed to the overall declined in Australia's PISA results since 2000 and those of several OECD countries. Unfortunately, there is no data to answer this directly. However, there is indirect evidence to suggest it is one factor among others.

The PISA results show increasing student dissatisfaction at school across OECD countries which may show up in reduced

effort and lower results. For example, the proportion of Australian students who feel unconnected with school increased fourfold from 8% to 32% between PISA 2003 and 2018. This was the 3rd largest increase in the OECD behind France and the Slovak Republic.

The fact that one-third of Australian students are dissatisfied with school is likely to manifest in low motivation and effort in tests that have no consequences for students because they don't ever see their results. The OECD says that the relationship between a feeling of belonging at school and performance in PISA is strong for those students with the least sense of belonging [OECD 2016, p. 122]. Students who feel they do not belong at school have significantly lower levels of achievement in PISA than those who do feel they belong.

Australia is not the only country with declining PISA results and increasing student dissatisfaction with school. PISA results for OECD countries have fallen since 2000 while the proportion of students who feel they don't belong at school increased threefold from 7% to 22%. Of 30 countries for which data is available, all experienced an increase in student dissatisfaction at school and PISA maths results fell in 21 of those countries.

The possibility that student effort on PISA has declined helps explain the contradiction between Australia's PISA and Year 12 results. Some 75-80 per cent of Australian students participating in PISA are in Year 10. While the PISA results for these students have declined since 2000, the results for students two years later in Year 12 have improved significantly.

The percentage of the estimated Year 12 population that completed Year 12 increased from 68% in 2001 to 79% in 2018 [Report on Government Services 2007, Table 3A.122 & 2020, Table 4A.70]. The completion rate for high SES students increased from 76% to 84% and from 62% to 75% for low SES students Year 12 assessments are high stakes in comparison to PISA even for less motivated students because these assessments have direct personal consequences for future education and work opportunities.

The important conclusion from the ACER and other studies of student motivation and effort is that the PISA results could be as much a measure of student effort as a measure of student learning. Therefore, they are not fully reliable as many assume and much caution is needed in interpreting the results and drawing strong policy conclusions.

The new results also raise the question as to the extent to which NAPLAN test results might also be affected by varying levels of effort and motivation by different groups of students. To date, no such research has been conducted. It should be on the research agenda for ACER and the new Australian Education Research Organisation to better inform the public and schools about the accuracy and reliability of NAPLAN.

9 November 2021

Trevor Cobbold National Convenor

SOS - Fighting for Equity in Education

https://saveourschools.com.au

https://twitter.com/SOSAust saveourschools690@gmail.com

Authorised by Sally Dennis, General Secretary, The State School Teachers' Union of W.A.
ABN 54 478 094 635 © 2026