

Home / Methodology / Limitations & challenges

Limitations & challenges

The limitations and challenges identified in PILNA 2021 are described below.

- Writing performance cannot be benchmarked as the writing proficiency scale and expected levels of performance for writing have yet to be developed. 2021 writing data will be re-analysed in future when these are available.
- PILNA cannot make causal claims between its contextual data and cognitive data.
- Comparisons between PILNA cycles for numeracy proficiency level eight are limited.
- Differences in school classifications exclude some information from the regional findings.

Separation of reading and writing scales

In 2021, reading and writing were separated for reporting purposes. This meant that different proficiency scales were needed for each domain.

Whilst reading and writing are both components of the PILNA literacy assessment and the skills involved in those areas are similar, both domains assess markedly different constructs. From a psychometric perspective, assessing reading and writing as separate constructs will strengthen the theoretical underpinning of the PILNA measurement model.

In addition to strengthening PILNA measurement, separate reading and writing assessments will enable more detailed evaluation of question quality, assessment quality, and performance in each area.

A separate reading scale also provides an opportunity to increase the robustness of comparisons between year four and year six performance. This in turn enables greater monitoring of growth in student performance over time.

No causal relationships in PILNA

The contextual component of PILNA was fully implemented for the first time in 2018. Contextual data are referenced significantly in the reports but only to provide descriptive context and indications of association with student learning outcomes.

When interpreting the results, care should be taken not to ascribe a causal relationship between contextual factors and student literacy and numeracy performance.

PILNA cannot, and does not, make any claims about contextual factors affecting student performance. Further analysis could usefully be conducted to identify relevant contributing factors, but, until that is done, care should be taken in the way the results are interpreted.

Comparing numeracy level eight between PILNA cycles

The splitting of Level 8 of the numeracy scale into 8a and 8b provides for a better description of what students know and can do at each of those levels.

At the same time, care should be taken in interpreting how the combined Level 8 results from 2012 and 2015 compare with 2018 and 2021.

The category was expanded through the incorporation of additional items in measurement, geometry, data and probability that were not part of the instruments in previous PILNA cycles.

Differences in school classifications

The categorisation of locality (urban, rural, remote, and very remote) is not included in the regional and small island state reports because of the differing definitions of locality in each of the PILNA countries. For example, one country identifies all its schools as rural.

Locality is, however, included as a subgroup in the individual country reports. Similarly, categorisation of school authority (government, non-government) is not included in the regional and small island states reports for the same reason.

School authority, where applicable, is included as a subgroup in the individual country reports.