

Reliability and validity of NAPLAN

The National Assessment Program—Literacy and Numeracy (NAPLAN) provides an assessment of the literacy and numeracy skills of students in Years 3, 5, 7 and 9. NAPLAN has a number of purposes including reporting national and State and Territory achievements in literacy and numeracy as well as providing accurate and reliable measures of student and school performance.

Processes have been put in place to ensure that NAPLAN is a valid and reliable measurement of students' literacy and numeracy skills. These processes are outlined below.

Development of NAPLAN tests

The tests are developed using the nationally agreed Statements of Learning that reflect the core elements of the curriculum documents used in the different States and Territories. The tests will reflect the Australian Curriculum in the future.

Tests are developed in five domains: Reading, Writing, Spelling, Grammar and Punctuation, and Numeracy. Specialist item writers are contracted to develop test questions. When constructing test questions, writers must meet nationally agreed specifications including curriculum coverage, spread of item difficulties, length of tests, question types (multiple choice or open response) and the need to have test questions that are common to adjacent year levels.

Quality assurance

Proposed test questions are provided to all State and Territory government and non-government education authorities for review. Assessment and curriculum specialists and classroom teachers, as well as specialists in Indigenous education, English as a second language, disabilities and inclusive education, review the items to ensure that the test questions are accessible to all students. This process ensures that only those questions that meet stringent criteria proceed to trial.

NAPLAN trials

Samples of students from all jurisdictions participate in NAPLAN test trialling. The trial checks whether sets of items that have been developed are suitable for inclusion in the NAPLAN tests, including whether there are sufficient items to cover the range of student abilities at each year level. Items are selected for the final tests using data from the trials, along with professional judgments from educational measurement, test construction and curriculum experts from all jurisdictions.

Expert advice

Educational measurement experts provide advice on test development and endorse the tests before they are finalised. These experts also provide advice and endorsement of key technical methods and specifications for reporting, equating and standards.

Common scales

In order to enable the comparison between years of testing, the tests have to be placed on common scales. The use of common scales covering Years 3, 5, 7 and 9 for each domain assessed allows for an individual student's achievement to be mapped as he or she progresses through schooling.

Equating between test years

As no two tests can have exactly the same level of difficulty, a rigorous equating process is carried out so that the difficulty of the current year's tests can be adjusted, if need be, to the same level of difficulty as the previous year's.

Equating tests are developed and administered so that current NAPLAN tests can be put on the same scale as previous years. The process involves a sample of students from each year, covering all States and Territories and school sectors, sitting the secure equating tests as well as the current year's tests. Using a combination of methods, the equating tests and current year's tests are placed on the same scales as previous years' tests.

Equating tests are delivered by specially trained independent test administrators. This ensures that the security of the equating tests can be preserved.

The equating process for NAPLAN was developed following expert advice. Great care is taken to provide a high level of assurance as to the reliability of comparisons between years. The equating process enables us to say with confidence that any test difference has been taken into account before making statements about one year's results compared to the next.

Reliability

While no test is one hundred percent reliable, there are well-established methods for estimating the reliability of tests. These methods indicate that the reliability of NAPLAN tests is high and that they can be used with confidence and are fit for purpose.

The rigorous processes that are carried out during the development of NAPLAN each year ensure that the results are reliable and comparable between years.

The Australian Council for Educational Research (ACER) advises that NAPLAN represents 'world's best practice in the ability to measure student progress'. The detailed report by ACER (Reporting and Comparing School Performance: 2008) is available on the Ministerial Council for Education, Early Childhood Development and Youth Affairs website (see www.mceecdya.edu.au).