

NAPLAN Review Interim Report

November 2019

**Barry McGaw
William Loudon
Claire Wyatt-Smith**

Contents

Purposes of standardised testing	1
Purposes of NAPLAN	1
Current concerns about NAPLAN	2
Deficiencies in tests	2
Stakes too high	2
Curriculum being narrowed	3
Results too delayed to be useful	4
Initial thinking about changes to deal with the concerns	4
Change the assessments	4
Change the timing	4
Change the content	4
Lowering exposure of results	5
Reducing display on My School	5
Removing data from My School	5
Making NAPLAN sample not census assessment	6
Making NAPLAN sample survey with opt-in	6
Next steps	6

We have just commenced our work on the review of the National Assessment Program: Literacy and Numeracy (NAPLAN) so this Interim Report is based on limited consultations and limited detailed work in relation to the Terms of Reference for our review. The starting proposition is that the agency of the profession in NAPLAN is essential if the test is to achieve its intended purposes. A related proposition is that teachers' professional judgement is the foundation of system approaches to promoting student growth and monitoring student learning and achievement over time. A third is that some kind of external check, either with standardised assessment or through consensus moderation of teachers' assessments by teachers from different schools, is necessary to ensure that assessments are comparable across schools.

We met in Sydney on Wednesday 2 September with those members of the Secretariat who are based in the NSW Department of Education. We then met with Ministers, representatives of government and non-government school authorities, regulatory authorities, teacher organisations and other experts in Melbourne on Tuesday 22 October, Brisbane on Wednesday 23 October, Sydney on Thursday 24 October and Canberra on Friday 25 October. In that four-day period, we held approximately 30 meetings and met with around 60 individuals.

We will continue to work in the remainder of this year. In the new year, we will engage in more extensive consultations, receive and review submissions and examine the administration of the NAPLAN tests and technical aspects of the scoring and scaling of the results.

In this Interim Report, we set out the major concerns about NAPLAN that we have heard or already knew about from our own work and offer some preliminary thinking about strategies for reducing or eliminating them.

Purposes of standardised testing

The primary purpose of standardised testing is to assess all students under the same conditions. External examinations at the end of Year 12 are standardised assessments as were the external examinations held in earlier times at the end of primary education and in mid-secondary education. NAPLAN tests are standardised. So are ACER's Progressive Achievement Tests (PAT), which schools may choose to use to monitor students' progress.

Purposes of NAPLAN

In our final report we intend to provide in an appendix a brief summary of the development of NAPLAN. At this point, we note that, over time, the purposes of NAPLAN and the uses to which it is put have increased. Some states and territories had literacy and numeracy testing programs as far back as the 1980s. One clear early purpose was to provide parents with assessments of their children's development in literacy and numeracy that were informed by a broader perspective than that of their children's individual schools.

The full data sets provided a picture of the overall development of students in the states and territories. In addition, the data could be analysed at the school level and, for the first time, there was the opportunity to use measures of students' socio-educational backgrounds to compare schools with similar students and to use the higher performing schools as sources of best practice for lower performing schools with similar students. Over time, then, system and school accountability and school and system improvement were added as purposes.

Analyses at the school level were shared with schools but not made public until Victoria published school-level measures in the years prior to the introduction of the My School website that made school-level results public for virtually all schools in Australia. With data over time available it became possible to shift the focus from status to improvement.

Current concerns about NAPLAN

Deficiencies in tests

When all students take the same relatively short tests, the tests will contain quite a few items that are so easy they will provide little information about the level of achievement of high-performing students and quite a few that are so difficult that they will provide little information about the level of achievement of low-performing students. Prior experimental work by ACARA showed that this should be so. We expect to have access to analyses of the performances of students in NAPLAN online 2019 to check this claim.

The assessment of writing, however, is more problematic. The move to NAPLAN Online has raised issues concerning the 'mode' of test taking. We expect to have access to analyses of the assessments of students' writing using pencil and paper or composing online using new technology. We will consider the utility of online functions, including those that could broaden writing stimuli for engaging students and keyboarding and word processing functions, addressing age appropriate online writing assessment. We will examine the effect of current time restrictions for writing in the online mode.

Stakes too high

The stakes in NAPLAN were raised once systems began to use the results for comparisons among schools. They are raised further when the results were made available to parents, and the community more generally, first through a Victorian Department of Education website and then for all schools in Australia through the My School website from 2010.

The website shows a school's results in comparison with national means. Beyond that, comparisons are made only among schools

with students from similar socio-educational backgrounds. These comparisons have revealed substantial variations in students' NAPLAN results among schools with similar students. Those variations challenge low-performing schools to recognise that more could be achieved with their students since other schools in similar circumstances are doing much better. The comparisons have also revealed that some schools, doing well when judged in comparison with state and national means, were actually doing relatively poorly when compared only with schools working with similar students. Those schools were coasting.

Once trend data became available comparisons of student growth were made among schools with students from similar socio-educational backgrounds but also between each school and a synthetic school constructed with the results of all students, appropriately weighted, with the same starting points as those of the students in the school.

While the My School website does not provide raw league tables that take no account of the differences in students' backgrounds, data can be extracted from the site for such comparisons. Major newspapers extracted data for this purpose in the early years though, more recently, most have focused on the information on student gains and reported on schools achieving the highest gains. In small communities, however, where there is a small number of schools, NAPLAN data from the My School website can be used to make unfair comparisons across the schools with no regard to differences in their students' backgrounds.

There are also some problems with the comparability of schools within a comparison group judged to have students from similar socio-education backgrounds. We will review this matter more fully in the new year for our final report, but we note here a particular case. It involves the inclusion of academically selective schools. For them, their students' socio-educational advantage has much less impact on the collective educational achievement of their students than does the inclusion of only high performing students. Comparisons with these schools, taking account of only students' socio-educational advantage, are inherently unfair.

The public availability of comparisons among schools on the My School website was intended to provide parents with information on school performance. Arming parents with this information has increased the stakes in NAPLAN results for schools. The stakes remain formally low for students since they have little direct impact on their further education beyond the possibility of support to deal with any deficiencies, except in any cases where NAPLAN results may be used as a basis for selection for entry to some secondary schools. We need to gather more evidence on this matter.

Curriculum being narrowed

NAPLAN has the potential to narrow the curriculum and thus students' learning in several ways. Teachers can narrow their teaching to only those aspects of the curriculum that they believe are tested by NAPLAN. This can reduce the curriculum, particularly in primary schools, to a focus on literacy/English and numeracy/mathematics, denying the students access to the rich array offered through other subjects.

Even if the teaching is not narrowed to the perceived test content, excessive time could be spent on taking prototype tests as rehearsals for the actual NAPLAN tests. Some time is needed to ensure students are familiar with the test form but, beyond that, more time spent in test rehearsal carries an important opportunity cost by denying time for richer learning.

Behind any tendency to limit students' experiences to a narrow literacy and numeracy curriculum lies a recognition that some important conclusions about a school will be based on narrow literacy and numeracy criteria. Supporters would claim that these are foundational competences with which all students need to be secure. Critics do not see expansion of the range of tests as a solution since they would judge many of the goals of schools still to be not well captured.

Results too delayed to be useful

The current NAPLAN schedule of testing in May and providing results to parents, students and schools in August to September, depending on the state or territory test administration authority, offers little opportunity for schools to use the results formatively to plan specific learning opportunities for either groups or individual students on the basis of achievements and/or deficiencies to date. That timeline has been shortened somewhat in the ACT where all students now take the test online and student-level results are provided to schools in July.

The move to NAPLAN Online would permit much speedier return of results for the components scored by computer if delivery were not delayed until the results of the writing test were available.

Initial thinking about changes to deal with the concerns

In the light of these concerns about NAPLAN, we offer some preliminary thoughts about changes to the assessment and reporting regimes.

Change the assessments

Change the timing

One possible change would be to shift the administration of NAPLAN from May to late February or early March. Results for the components tested online could be provided within a couple of days, if not even the next day. Early testing in the year would restrict the opportunity to waste time teaching to the test.

The speedier delivery of the results would improve their usefulness, but that could be achieved whenever the tests were delivered. Shifting the tests to early in the year, combined with speedy delivery of results, would make NAPLAN a measure of teachers' and students' starting points for the year. It could liberate NAPLAN to play a formative rather than a summative assessment role and to inform decisions about future curriculum and teaching choices, not judgements about past ones.

It is, of course, possible that start-of-year assessments would be seen as summative assessment of the end of the previous year. That argument is potentially weakened by the impact of declines in student performance over the summer vacation and the tendency for class groups to be formed with different mixes of students in each new year of schooling.

Assessment of the starting points for the year could give school systems the opportunity to provide additional resources to schools in most need of additional support.

Coupled with a move to testing earlier in the year could be a reconsideration of the school years in which the tests are administered. The present schedule is Years 3, 5, 7 and 9. Would Year 3 be too early if the tests were at the beginning of the school year? On the other hand, would waiting until the beginning of Year 4 be too late, given the importance for students' academic self-concept of becoming secure readers early in their school lives?

For separate secondary schools, would it be better to obtain information on incoming students early in Year 7 or to wait until Year 8, when the students are potentially more secure in their new environment?

A further consideration could be whether to administer NAPLAN only in Years 3 and 7. That would reinforce its potential formative role but would deny the possibility of measures of growth achieved by schools for those students remaining in the same school from Years 3 to 5 (or Years 4 to 6) and Years 7 to 9 (or Years 8 to 10).

Change the content

A review of NAPLAN offers an opportunity to review the content and coverage of the tests. We need to review the assessments of reading, language conventions and numeracy before offering any specific judgements or suggestions about them. We also plan to consider whether NAPLAN might be expanded to include any of the General Capabilities in the Australian Curriculum and connections to Literacy Learning Progressions.

We are, however, already in a position to offer some judgements of the assessment of writing in NAPLAN and to offer some suggestions for change. The NAPLAN writing task is designed to assess accurate, fluent and purposeful writing of either a narrative or persuasive text, with performance assessed against ten criteria: audience, text structure, ideas, character and setting (for narrative text)/persuasive devices (for persuasive text), vocabulary, cohesion, paragraphing, sentence structure, punctuation and spelling.

The limited choice of prompts (even with separate Years 3, 5, 7 and 9 prompts) and text types or forms of writing that students are asked to produce should be reconsidered. In its current form the writing test requirements are reported to lead to formulaic pieces that reflect student attempts to reproduce rehearsed writing prescriptions.

The nature and function of the NAPLAN writing rubric should be reconsidered, examining scope of the stated criteria, the implications of the numeric scores, and technical issues of scoring including dependencies among criteria, especially in relation to adjacent year levels. There is potential benefit from investigating the difficulty and ease with which scorers can separate the criteria for scoring purposes, discriminating among them and addressing the specified features within each criterion.

We think that careful research into the NAPLAN writing task and the assessment of student responses should be undertaken and that this could best be done by assessing writing for samples of students, not the full Years 3, 5, 7 and 9 cohorts. Once the research has been completed and its conclusions reviewed, consideration could be given to returning to the assessment of all students. The research on NAPLAN writing should address:

- handwriting, word processing and keyboarding as they relate to student performance in writing and student development over time
- broadening the types of writing assessed, recognising that, to date, NAPLAN writing has assessed only narrative writing and persuasive writing but not informative writing which is also included in the Australian Curriculum: English
- the number and type of prompts, including the use of multi-modal prompts if they are presented online
- the time allocation required to allow students to undertake drafting and editing to produce final copy

- the nature and function of scoring criteria
- alignment of the writing test with the General Capabilities and Literacy Progressions in the Australian Curriculum, since those links appear not to be made clear to the teaching profession.

If there were significant changes in the content of NAPLAN tests, it might also make sense to begin a new time-series for the tests. New tests, or tests in new areas, would provide opportunities to design tests that maximise the opportunities of online assessment. Within two or three cycles of testing, a new set of time-series and achievement gain calculations would be available.

Lowering exposure of results

There appear to us, at this stage, to be more concerns about the impact of uses of NAPLAN data than about the tests themselves.

Reducing display on My School

One way to reduce the impact of public disclosure of NAPLAN results would be simply to reduce the extent of use of NAPLAN on the My School website, as Bill Loudon proposed earlier in the 2019 NAPLAN Reporting Review prepared for Education Council. It could, for example, be reduced to a single page with information on status in the current year, trends over years for the school and change between two school years of testing for students in the same school in both years.

Removing data from My School

A more radical proposal would be to remove NAPLAN data from the My School site. Since the site was created to report primarily on NAPLAN results, would the site outlive the removal of NAPLAN results? What would be the impact of the loss of public reporting on other aspects of the schools that have been introduced through developments of the My School site, perhaps most particularly the data on school income from all sources and expenditure?

The loss of public revelations of schools' NAPLAN results on a single site could be offset by a requirement that each school publish particular aspects of its results prescribed in some template. It could also be offset by release of anonymised school and student data in a form that would facilitate research into aspects of schooling that the My School website does not enable. Results could still be linked to the measure of socio-educational advantage, but they could also be linked to results of students participating in sample surveys, both international such as PISA, TIMSS and PIRLS and the national surveys of Civics and Citizenship, ICT Literacy and primary science.

These educational data could also be linked to other data, such as health data, as Western Australia has already achieved.

Making NAPLAN sample not census assessment

More radical still would be to assess only a sample of students in NAPLAN and not assess the whole cohort of students in the particular year levels assessed. Data from a sample assessment would permit monitoring of the system over time and, through appropriately structured random samples, also permit monitoring of subpopulations of interest, such as states and territories, males and females and Aboriginal and Torres Strait Islander students.

What a sample survey would not provide is student data with which to report to parents or school-level data with which systems could identify schools that are underperforming taking account of their context, including schools that are doing well yet coasting.

Without data on all students and so all schools, individual schools would not be able to see themselves in a larger framework and make informed judgement of themselves.

It might be possible to structure the samples by drawing them 'without replacement' of schools in prior samples for a couple of years so that all schools would be sampled once every three or four years. That would provide something of a repeating census over time but not annually.

Making NAPLAN sample survey with opt-in

A compromise approach would be to make NAPLAN a sample survey but to allow schools not sampled to opt in. System-level data and analysis would come exclusively from the random sample but the other schools opting in would receive results that were calibrated on the national scales and so enable them to see themselves in the national and state and territory context.

Similarly, an opt-in provision for students would be possible, with parents in any school able to seek to have their children assessed and to receive individual reports calibrated to national scales.

An important question would be whether an opt-in provision would become coercive, putting pressure on schools to opt in and not stay out as having not been included in the sample.

Next steps

We are conscious that the Interim Report offers only provisional thinking about the issues before us. We would not want our report to suggest that we have ourselves ruled anything in or out at this stage nor would we want Ministers or officials to rule things in or out until we have provided our final report.

NAPLAN Review Interim Report

We acknowledge the homelands of all Aboriginal people and pay our respect to Country.