



22 March 2019

Submission to the NAPLAN Reporting Review 2019

Peter Goss and Julie Sonnemann

Summary

We welcome the opportunity to present our views to the NAPLAN 2019 Reporting Review, commissioned by The Education Council of the Council of Australian Governments. Our submission addresses the terms of reference on how to improve presentation of NAPLAN, but also draws attention to the broader debate on NAPLAN.

We argue that NAPLAN should not be scrapped, nor moved to sample testing; it is a vital tool for adaptive education systems to monitor student performance and improve government support over time. Governments and other system leaders rely on standardised testing to understand which schools are struggling or thriving, and what interventions work well and should be expanded.

We note that current NAPLAN reporting works reasonably well as a monitoring tool, but much less effectively for evaluating what works or as a tool to inform parents on their school choice.

We also highlight two things NAPLAN should not be used for. First, NAPLAN reporting should not aim to stimulate competition between schools; there is little evidence this approach will improve teaching in Australia.

Second, NAPLAN should not be expected to support teachers as a diagnostic tool in the classroom, even with the improvements to NAPLAN online. There are benefits in keeping separate the standardised assessments intended for monitoring and accountability from the classroom assessments that teachers use regularly to improve what they do.

We recommend the following changes to NAPLAN reporting:

- Raise the national minimum standard or stop reporting it.
- Report NAPLAN learning progress using a measure that is comparable across students from different starting points.
- Improve the presentation of results on My School, in particular making it easier for parents to access student gain results and school trends over time.
- Strengthen the Annual NAPLAN report by including more analysis on learning gain and by contextualising comparisons among states and across geolocations.
- Support third-party reporting that uses NAPLAN data by:
 - Simplifying access to the unit record data; and
 - Improving linkages to other data.

Finally, we suggest that the Education Council consider:

- Expanding NAPLAN data to cover general capabilities; and
- Expanding school-level data (especially on My School) to include information about educational practices.

Table of contents

Summary	1
1. Putting this review in context.....	3
2. NAPLAN is a critical link in the data ecosystem.....	7
3. How to improve NAPLAN as a monitoring tool	9
4. How better access to NAPLAN data would improve evaluation.	12
5. How better reporting would support school choice	13
Bibliography.....	14

1. Putting this review in context

1.1. Context of current review

NAPLAN is a national asset. With 12 years of data, gathered four times during the course of every student's schooling, it provides vital insight into how schools and students are performing. Without NAPLAN, we would know much less about the outcomes and effectiveness of school education in Australia.¹

Yet there are legitimate questions about whether NAPLAN has delivered what it promised, and the negative impacts it can have on students and schools. These questions go beyond the inevitable limitation that the desirable outcomes from schooling are broader than any standardised test can cover.² On balance, we believe that NAPLAN and the way it is reported do more good than harm. But improvements are possible and desirable.

This review is therefore welcome. This is especially true because the review occurs in a context where some stakeholders are asking for bigger changes to NAPLAN than just reporting. Some are calling for NAPLAN to be scrapped, or reduced to sample testing. Before responding more directly to the context of the review, we put forward our views on these two broader issues.

¹ Goss (2018).

² However, Year 9 NAPLAN results help predict which students will leave school before Year 12, and employment outcomes for early school leavers. ABS (2014).

³ See, for example, <https://www.theeducatoronline.com/au/news/scrapped-naplan-expert-urges/246353>.

⁴ See Goss (2017), p. 26.

Don't scrap NAPLAN

Some stakeholders argue that NAPLAN should be scrapped in favour of other diagnostic assessments used regularly by teachers to improve their own teaching practice.³ The argument is that this would remove the perverse incentives created by the perceived 'high stakes' associated with NAPLAN.

On some level, this argument has merit: our education systems would be more adaptive if teachers and schools were better able to track the progress of their students in ways that directly inform their teaching in the classroom.⁴ NAPLAN is the wrong tool for this purpose. The assessment within NAPLAN is too narrow and too infrequent to enable targeted teaching in the classroom.⁵

But NAPLAN is a standardised test that is inherently linked with public and political accountability and monitoring – essential in a public education system. It helps governments monitor school performance and understand what works. If NAPLAN were removed, teacher-generated data would inevitably become used for government monitoring and accountability.⁶ This would harm the trust that is so vital in teacher-generated data.

⁵ See Goss et al. (2015), p. 13.

⁶ Without NAPLAN, our judgement is that the desire for top-down accountability would overwhelm the legitimate argument to keep teacher-generated data focused on improvement and collective professional responsibility.

Those who want NAPLAN scrapped should be careful what they wish for. There are benefits in keeping separate the standardised assessments governments want for monitoring and accountability, and the classroom assessments that teachers use to improve what they do. Dangers can arise when the two goals get blurred.

Australia is not ready to move to sample testing

Others have argued that we could keep the benefits and reduce the downsides of NAPLAN if it were a sample test.⁷ The problem is that teacher judgment is not sufficiently rigorous in all schools, nor linked to common standards.⁸ Without NAPLAN, the risk is that even more schools and students would fall through the cracks.

All educators (indeed, all professionals) need to verify their own judgments against independent and objective data. NAPLAN is not inherently necessary for such independent verification; but there isn't currently anything ready to replace it in Australia at scale.

⁷ Piccoli et al. (2019).

⁸ See Goss et al. (2015), pp. 11-12.

⁹ This is separate from the strengths and limitations of how NAPLAN is used.

¹⁰ See Goss et al. (2018).

¹¹ See Goss et al. (2016).

1.2. Strengths and limitations of NAPLAN as a test⁹

In our view, the key strengths of NAPLAN as a test are its:

- **National consistency**, which enables comparisons across schools, sectors, and states;¹⁰
- **Contextual information**, which enables like-for-like comparisons across schools and student groups;¹¹ and
- **Common scale** across year levels, which enables analysis of student learning growth.¹²

NAPLAN's key limitations as a test are that:

- **Measurement error** makes NAPLAN data much less useful for individual students or small schools;¹³
- The **NAPLAN curve** makes it hard to interpret student learning growth;¹⁴
- The **National Minimum Standard** is set too low;¹⁵ and
- **Variable participation rates** can make it hard to compare groups of students or schools.

¹² See Goss et al. (2016) and Goss et al. (2018).

¹³ See Wu (2010).

¹⁴ Goss et al. (2016) proposes an 'Equivalent Year Level' metric to account for the curve. This was updated in Goss et al. (2018).

¹⁵ See Goss et al. (2016), pp. 23-24.

1.3. NAPLAN reporting is working better for monitoring; but less so for evaluation and parent choice

This review is about NAPLAN *reporting*. While the issues paper focuses heavily on the My School website, NAPLAN reporting needs to be considered broadly,¹⁶ because each reporting channel raises specific considerations and has different goals.

NAPLAN reporting takes place through both public and private channels. Public reporting channels include the My School website; the NAPLAN annual report, and its online version; and third-party analysis of NAPLAN, such as media stories or reports by organisations such as Grattan Institute. Private reporting channels include the provision of NAPLAN data to education systems, schools, and parents.

There are also different goals for reporting: monitoring; evaluation; and parental choice (see Box 1). The current reporting model works much less effectively for evaluation and parental choice than for monitoring. The NAPLAN reporting issues associated with monitoring, evaluation and parental choice are discussed in more detail in Chapters 3, 4 and 5.

Box 1. Reporting on NAPLAN has three main purposes

Monitoring. School and system leaders use NAPLAN to monitor the achievement and progress of their students.

The key monitoring question is “what do the data tell us about a specific group of students or schools?”¹⁷

Evaluation. Policy makers and researchers also use NAPLAN to better understand interventions that lift student performance, inform system-wide policies, and target support to schools. The point is to use current data to improve future performance.

The key evaluation question is “what do the data imply about the effectiveness of schools (or systems, educational interventions, etc)?”

Parental choice. Parents use NAPLAN to inform their choice of school.

The key question in parental choice is “will this school educate my child effectively?”¹⁸

¹⁶ By contrast, the issues paper does not mention the NAPLAN annual report or its online version at <https://reports.acara.edu.au/Home/Results>.

¹⁷ The need to monitor all students and schools for accountability purposes is a key reason why NAPLAN should not become a sample test.

¹⁸ This applies both when parents are choosing a school for their child, and when they are deciding whether to keep their child at a school.

1.4. How NAPLAN should not be used

There are two purposes for which NAPLAN is not the best tool.

Competition

NAPLAN reporting should not aim to stimulate competition between schools.

The My School website gives families information on how schools perform in NAPLAN. In theory, this information could help stimulate competition between schools. In practice, this has not happened.

Relying on school markets is not the best way to improve student learning. In Australia, families generally don't move to high-performing schools nor leave low-performing ones.¹⁹

Targeted teaching

NAPLAN should not be expected to support teachers as a diagnostic tool for individual students in the classroom. Nor can it identify what each student knows so that the teacher can target their teaching to what the student needs to learn next.²⁰

NAPLAN assesses two years' worth of learning in each subject area through about 35-40 questions, most of which are multiple choice. It is not clear that all schools recognise the high level of measurement error in individual students' NAPLAN scores. By

chance, a student's score may be out by more than half a year's learning. The error in measuring student growth is higher still.

In addition, NAPLAN tests are designed to have broad coverage, not to diagnose in detail what individual students are ready to learn next or the underlying source of any difficulties they face. Yet that is what targeted teaching needs.

Moving NAPLAN to online adaptive testing will make it more accurate and return results sooner, but not address all the issues outlined above. While it has many benefits, NAPLAN is not sufficient to comprehensively assess individual students' learning or track their progress.

¹⁹ Jensen et al. (2013).

²⁰ Goss et al. (2015), p. 13-15.

2. NAPLAN is a critical link in the data ecosystem

School education in Australia has many bright spots, but we do not have a system of excellence or an *adaptive education system* that identifies excellence and systematically spreads it.²¹

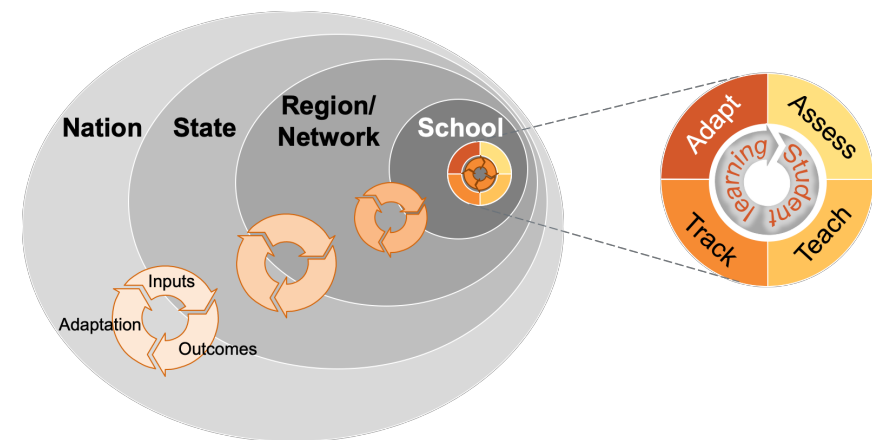
Data about student achievement and progress is an essential component of an adaptive education system. But no single assessment can possibly provide the data required to inform educational practice. ‘Small data’, classroom assessments done by teachers on a regular basis, are essential to guide the teaching and learning process at a local level. Done well, such data is more *relevant* than a standardised test (what we call ‘big data’) can ever be. But it is also less *rigorous* than the big data generated by standardised tests like NAPLAN, at least in enabling comparisons over time and across schools.

Data about student learning progress needs to be used at multiple ‘levels’ within education systems. Figure 1 shows adaptive improvement as a series of nested feedback loops. For feedback to work, educators must look at the practices they are currently using (‘inputs’), the impact on student learning (‘outcomes’), and have a systematic adaptation process for deciding what to keep doing and what to stop (‘adaptation’).

Different assessment tools are needed to complete the feedback loop at different levels of education systems. NAPLAN is the wrong tool for targeted teaching (the feedback loop within

schools).²² But it is highly valuable for the improvement loops that are required across schools, across regions, at a state government level and across states.

Figure 1: An adaptive education system needs feedback loops at multiple levels.



Source: Goss (2017).

²¹ (P Goss, 2017).

²² In our view, online NAPLAN will not change this. It will make the data more accurate (particularly for high- and low-achieving students), but the data will still

be too infrequent and too narrowly defined to be the main input into targeted teaching. Online NAPLAN data should, however, be welcomed as a way to independently verify internal school assessments of literacy and numeracy.

Australia's challenge is that, at times, NAPLAN can play too dominant a role, like a heavy weight that unbalances a set of kitchen scales. The best way to get to balance is not to throw out NAPLAN, but to strengthen the counter-balancing forms of assessment, particularly small data in the hands of teachers.

Australia's national data are also poorly balanced between inputs and outcomes. As the diagram shows, adaptive improvement needs data about what is actually being done in schools, as well as what students have learned. Schools and education systems can't benchmark themselves to better outcomes without linking the results to the actions that contributed to them.

Our main focus as a nation should not be on tinkering with NAPLAN but in trying to define a more rounded data ecosystem that incorporates NAPLAN but takes us beyond it, e.g. by incorporating measures of quality teaching. If it were part of a broader ecosystem of data, the current downsides of NAPLAN would be greatly reduced.

In the meantime, NAPLAN reporting should certainly be improved. But developing an effective data ecosystem to drive continuous improvement is more about the overall balance of data – especially improving the quality of data gathered through day-to-day teacher assessments, or gathering more systematic information about practice – than it is about tweaking NAPLAN.

3. How to improve NAPLAN as a monitoring tool

3.1. Monitoring literacy and numeracy at a system level

Australia spends about \$30 billion each year on primary school education. Yet nearly 3 in 10 Year 7 students lack the core reading skills they need to succeed in secondary school.²³ If nothing else, the need for political accountability makes it essential to monitor student outcomes using data that can be compared across schools and jurisdictions.

The key monitoring question is “what do the data tell us about the performance of a group of students or schools?” Performance means achievement (what do students know) and progress (how has this changed during the course of students’ schooling).

State governments and system leaders should use the answers to this question to set directions and inform policies. They should also monitor student performance in individual schools or groups of schools (e.g. a region) to identify where performance is strong and where extra support is needed.

²³ Lamb et al. (2015).

²⁴ Without moderation, teachers tend to grade in highly variable ways. See Connolly et al. (2012); Harlen (2005a); and Harlen (2005b).

²⁵ See Goss et al. (2016), Recommendation 2b, based on analysis that shows that a Year 9 student reading at the NMS is below the typical Year 5 student.

²⁶ NAPLAN gain scores are not directly comparable across students from different backgrounds because “students who start with lower NAPLAN scores tend to make greater gains over time than those who start with higher NAPLAN scores.” ACARA (2015), p. 5. Goss et al. (2016), Figure 2 shows how a face-

Standardised tests are not the only way to ensure that the data used for monitoring are comparable. But the main alternative – carefully moderated teacher-assessments against common learning standards – could be even more onerous.²⁴ And there is little reason to invest in an alternative when NAPLAN is ideally positioned to continue to provide raw data for monitoring.

Four changes would make reporting more effective for monitoring:

- Raise the **national minimum standard** or remove it entirely.²⁵
- Report **learning progress** using a measure – such as our proposed Years of Learning Progress metric – that is comparable across students from different starting points;²⁶
- Strengthen the **Annual NAPLAN report** by including a wider range of analysis on learning gain;²⁷ and
- **Contextualise state-by-state comparisons** in the Annual NAPLAN report, as well as comparisons across geolocation.

value interpretation of gain scores can suggest students are catching up when they are actually falling further behind.

²⁷ Student learning progress is the best measure of the effectiveness of schools and systems (see, e.g. Jensen (2010); Goss et al. (2015); and Goss et al. (2016)). Yet only about 10 per cent of the 2017 NAPLAN annual report (ACARA (2017)) was devoted to analysis of student gain (40 pages out of 365). The ACARA website reports.acara.edu.au has the same limitation because it presents the same data.

This last point is a subtle one and requires explanation.

The Issues paper says that “reporting on the outcomes of schooling should use data that is valid, reliable, and *contextualised*” (emphasis added). This is done carefully on the My School website, using ICSEA (the Index of Community Socio-Educational Advantage) as the basis for identifying comparable schools.

But the Annual NAPLAN report fails to acknowledge the socio-economic differences among states and territories.²⁸ Thus, for example, the ACT is routinely shown as having high levels of achievement, even though its performance on a like-for-like (i.e. contextualised) basis is relatively weak.²⁹ Meanwhile, the Northern Territory is shown as having low levels of achievement, without contextualising its much lower levels of socio-economic status.

The Grattan Institute report “Measuring Student Progress” was written in part to address this lack of contextualisation across regions, states, and sectors.³⁰ ACARA should routinely report contextualised comparisons, to enable politicians and the public to compare literacy and numeracy levels in a way that acknowledges important differences.³¹

²⁸ The Annual NAPLAN report also fails to acknowledge socio-economic differences between students in different geolocations.

²⁹ See, for example, Goss et al. (2018).

³⁰ Goss et al. (2018).

3.2. Informing school leaders

NAPLAN provides a range of valuable information to school leaders as they monitor academic performance in their schools and decide where to focus their scarce time and resources. But NAPLAN is only one source of information among many, and there are risks in focusing too strongly on any one indicator.

NAPLAN’s unique value for school leaders is that the data are comparable across year levels and across schools – and that the data about other schools is available.³² While teacher-generated assessment data are more *relevant* to day-to-day teaching and learning, NAPLAN data are more *rigorous* – at least when the sample size is large enough.

This rigor brings real value. Comparing average NAPLAN achievement to similar schools gives an indication of where students are performing above expectations and where they might be expected to do better. Comparing progress against similar schools (or similar students) is a more direct way of identifying where the school is adding the most value, and where it might lift its game. Identifying trends over time shows where improvement efforts are working, or where they are not.

This comparability can come with a cost. If schools focus on improving NAPLAN scores as an end in themselves, they may

³¹ The unadjusted results should continue to be published alongside contextualised results, because absolute levels of literacy and numeracy matter as well as the ‘value-add’ results.

³² Other standardised tests (e.g., the Progressive Achievement Tests (PAT) offered by the Australian Council of Educational Research) generate comparable data, but data about other schools’ performance is generally not available.

focus less on other elements of schooling that are equally important but less visibly measured.

Viewing NAPLAN as a 'high stakes' test may cause unnecessary stress for teachers and students. A different way to use NAPLAN that should cause less stress is as one element of an ongoing discussion about achievement and progress levels within the school.

For example, a secondary school might analyse its data to identify that a substantial minority of its Year 7 students routinely score at NAPLAN band 4 in reading – the average performance of a Year 3 student. NAPLAN is not an accurate diagnostic test for each student. But the leaders of this hypothetical school should take the overall pattern of the data very seriously. First, they should introduce mechanisms to quickly diagnose the reading abilities of all incoming students. Second, they should think about how the timetable needs to be arranged to accommodate so many students who are still learning to read, at the same time as providing adequate challenges for those students whose reading is at or above level.³³

NAPLAN reporting should make it as easy as possible for school leaders to see trends over time, both in their school and others;³⁴ and to make it easier for them to identify patterns within the data.

³³ Thanks to Ingrid Sealey, previously of Fogarty EDvance, for suggesting this way of thinking about the use of NAPLAN data by a school.

³⁴ Trends over time are more reliable than data from a single year. This is particularly true for data about growth, which fluctuate greatly from year to year.

4. How better access to NAPLAN data would improve evaluation

Better evidence about what has (and hasn't) worked in the past will enable better decisions about what to do in the future. The key question is, "what do the data imply about the effectiveness of schools (or systems, educational interventions, etc)?"

NAPLAN is an important data source, but My School and the NAPLAN Annual report provide only a glimpse of the richness of the NAPLAN data sets. More detailed evaluation needs access to unit-record data (as we used in our 2016 report *Widening Gaps*) or the longitudinal school-level data set that underpins My School (which we used in our 2018 report *Measuring Student Progress*). Others have used similar datasets for their analyses.³⁵ We note three main areas where better access to NAPLAN data would strengthen the evidence base. In turn, this would strengthen third-party reporting about NAPLAN.

Simplify access to unit-record data and school-level data

It is hard to get access to the detailed NAPLAN data. This dramatically limits the ability of third-party researchers to use the potential power of NAPLAN data.³⁶ Access to unit-record data is particularly important for researchers such as Grattan Institute who wish to transform the data before analysis.³⁷

³⁵ See, for example, Bonnor and Shepherd (2016); and Joseph (2019).

³⁶ For example, is variation higher within or among schools? PISA data suggests that variation in student achievement is higher within schools (OECD (2016), p. 226). But what about variation in student progress?

³⁷ For example, calculating the Equivalent Year Level of a state's average NAPLAN score gives a different answer than reversing the order and averaging the Equivalent Year Level data of every student in the state.

Privacy considerations mean that detailed NAPLAN data must be managed carefully. But this challenge has been solved for other sensitive government data, whether through anonymised sample data³⁸ or 'locked rooms' where analysis uses a comprehensive data set but only aggregated data can be retained.³⁹

Improve linkages to other data

A national student identifier would make it easier to link NAPLAN data to other datasets.⁴⁰ This could help researchers better understand the links between early childhood education and schooling, or between school attendance and student progress.

Expand school-level data to look at educational practices

As Chapter 2 describes, continuous improvement relies on comparing educational outcomes with educational practice. Australia has much more data about the former than the latter. This makes it hard for researchers to identify what practices might actually be causing over- or under-performance among schools, regions or states.

³⁸ For example, the ABS Confidentialised Unit Record File (CURF) dataset.

³⁹ The ABS DataLab provides a virtual 'locked room', where the analysis is done on ABS computers and an ABS employee approves any output.

⁴⁰ As an example of the power of data linkage, the National Schools Resourcing Board used the Multi-Agency Data Integration Project (MADIP) dataset as part of its review of the SES score methodology.

5. How better reporting would support school choice

Australia's parents expect to have the right to choose the best school for their child. Various structures – notably the My School website – are designed to help parents make these choices.

But while school choice can benefit individual students, it can also reduce equity and school quality. The 'non-choosers' often end up in schools with higher concentrations of student disadvantage and lower levels of student achievement. The OECD advises that school choice should be carefully managed to avoid some of the detrimental impacts on equity.⁴¹

School choice can also unintended impacts on school quality. If parents choose schools using a very narrow set of indicators, it can push the system toward 'lowest-common-denominator' approaches, such as teaching to the NAPLAN test and narrowing the curriculum to focus more on the areas within NAPLAN.

Parents need better information on school performance

The My School website should be improved in three key ways to provide parents with better information when choosing schools.

First, it should provide broader information on the quality of teaching at the school, rather than focusing narrowly on outcomes data, i.e. NAPLAN test results. Effective teaching is a key

influence on student learning, yet little data is collected on it in Australia. By contrast, the UK system provides parents with public quality assured judgments on teaching quality, and school leadership, as well as student results.⁴²

Second, My School should give parents better information on school performance in developing students 'general capabilities', such as resilience, collaboration, communication skills and so on. This would need to be a long-term goal, because we don't yet know enough about how to measure or even teach these capabilities.

Third, the presentation of NAPLAN results on My School should be improved. For example, the student gain results should be easier to find and interpret.⁴³ And it should be easier for parents to observe school trends over 3-to-5 years, rather than just a given year, because NAPLAN results can fluctuate a lot from year to year.

⁴¹ OECD (2012).

⁴² The UK body responsible for inspecting schools, Ofsted, provides expert judgments to parents on a broader range of school effectiveness metrics. See Roberts et al. (2019).

⁴³ In particular, student gain relative to students with the same starting scores should be better highlighted to parents.

Bibliography

- ABS (2014). *Educational outcomes, experimental estimates, Tasmania 2006-13*.
<http://www.abs.gov.au/ausstats/abs@.nsf/mf/4261.6?OpenDocument>.
- ACARA. (2015). *My School fact sheet: Interpreting NAPLAN results*. Retrieved from
http://www.acara.edu.au/verve/_resources/Interpreting_NAPLAN_results_file.pdf.
- ACARA (2016). *'Interpreting NAPLAN results'*, from
http://www.acara.edu.au/verve/_resources/Interpreting_NAPLAN_results_file.pdf.
- ACARA (2017). *NAPLAN Achievement in Reading, Writing, Language Conventions and Numeracy: National Report for 2017*.
- ACT Auditor-General (2017). *Performance information in ACT public schools*. ACT Auditor-General.
http://www.audit.act.gov.au/__data/assets/pdf_file/0017/1180007/Report-No-4-of-2017-Performance-information-in-ACT-public-schools.pdf.
- Bonnor, C. and Shepherd, B. (2016). *Uneven Playing Field – The State of Australia's Schools*. Centre for Policy Development, Sydney.
<https://cpd.org.au/2016/05/unevenplayingfield/>
- Connolly, S et al. (2012). *'Moderation and consistency of teacher judgement: teachers' views'*, British Educational Research Journal, 38(4), p 593-614
- Goss, P. (2017). *Towards an adaptive education system in Australia*. Discussion Paper No. 2017-01. Grattan Institute.
<https://grattan.edu.au/report/towards-an-adaptive-education-system-in-australia/>.
- Goss et al. (2015). Goss, P., Hunter, J., Romanes, D. and Parsonage, H. *Targeted Teaching: How Better Use of Data can Improve Student Learning*. Report No. 2015-06. Grattan Institute. <https://grattan.edu.au/report/targeted-teaching-how-better-use-of-data-can-improve-student-learning/>.
- Goss, P. and Sonnemann, J. (2016a). *Widening gaps: what NAPLAN tells us about student progress*. Report No. 2016-03. Grattan Institute.
<http://grattan.edu.au/report/widening-gaps/>.
- (2016b). *Grattan submission to the PC Inquiry into the National Education Evidence Base*.
<https://grattan.edu.au/news/submission-to-the-productivity-commission-inquiry-into-the-national-education-evidence-base/>.

- Goss, P., Sonnemann, J., and Emslie, O. (2018). *Measuring student progress: A state-by-state report card*. Report No. 2018-14. Grattan Institute.
- Goss, P. (2018). "Five things we wouldn't know without NAPLAN." *The Conversation*. <https://theconversation.com/five-things-we-wouldnt-know-without-naplan-94286>.
- Harlen, W. (2005a). 'Teachers' summative practices and assessment for learning—tensions and synergies', *Curriculum Journal*, 16(2), p 207-223
- (2005b). 'Trusting teachers' judgement: Research evidence of the reliability and validity of teachers' assessment used for summative purposes', *Research Papers in Education*, 20(3), p 245-270
- Jensen, B. (2010). *Measuring what matters: Student progress*. Report No. 2010-01. Grattan Institute
- Jensen et al. (2013). Jensen, B., Weidmann, B. and Farmer, J. *The myth of markets in school education*. Report No. 2013-07. Grattan Institute.
- Joseph, B. (2019). *Overcoming the Odds: A study of Australia's top-performing disadvantaged schools*. Research Report 39. Centre for Independent Studies. <https://www.cis.org.au/publications/research-reports/overcoming-the-odds-a-study-of-australias-top-performing-disadvantaged-schools/>
- Klenowski, V. (2009). *Raising the stakes: the challenges for teacher assessment*, AARE International Education Research Conference - 2009 from <http://eprints.qut.edu.au/43916/>
- Lamb, S. et al. (2015). *Educational opportunity in Australia 2015: Who succeeds and who misses out*, Centre for International Research on Education Systems, Victoria University, for the Mitchell Institute, Melbourne: Mitchell Institute. <http://www.mitchellinstitute.org.au/wp-content/uploads/2015/11/Educational-opportunity-in-Australia-2015-Who-succeeds-and-who-misses-out-19Nov15.pdf>
- Lamb, S. (2017). *Government school performance in the ACT*. Centre for International Research on Education Systems. https://www.education.act.gov.au/__data/assets/pdf_file/0011/1070399/Government-School-Performance-in-the-ACT-Stephen-Lamb-Redacted.pdf.
- Macintosh, A. and Wilkinson, D. (2018). *Academic underperformance in ACT schools: An analysis of ACT school performance in NAPLAN over the period 2012 to 2016*. ANU Law School. https://law.anu.edu.au/sites/all/files/act_naplan_performance_2012-2016_august_2018.pdf.
- Picoli, A. and Sahlberg, P., (2019). *Gonski Institute for Education: Submission to the Education Council of the Council of Australian Governments (COAG) NAPLAN Reporting*

Review 2019. University of New South Wales.
https://education.arts.unsw.edu.au/media/EDUCFile/NAPLAN_submission_21_March_.pdf?mc_cid=6163e958f6&mc_eid=ec4dc0008d

Roberts et al. (2019). *School inspections in England, Ofsted, Commons Library Briefing*.
<https://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN07091>

OECD (2012). *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*, OECD Publishing.
<http://dx.doi.org/10.1787/9789264130852-en>

OECD (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA, OECD Publishing, Paris.
<http://dx.doi.org/10.1787/9789264266490-en>

Wu, M. (2010). 'Measurement, Sampling, and Equating Errors in Large-Scale Assessments', *Educational Measurement: Issues and Practice*, 29(4), p 15-27