

Reassessing assessment

Dr Sandy Heldsinger co-founder of Brightpath

'Cutting-edge': the project that could replace NAPLAN. So read the head line to a recent *Sydney Morning Herald* article. The article went on the report that senior education officials are working on an ambitious alternative to NAPLAN that will track every student's progress and use low-stakes classroom tests to check how well students grasp skills and concepts.

It's not really a surprise, is it?

Many readers will know of the tittering in the educational community about the fact that spell check autocorrects NAPLAN to NAPALM. I think it's safe to say that some of our colleagues could well have preferred to have been working with NAPLAM this year, rather than NAPLAN.

Technology glitches aside though, testing time generates some fervour in the media and the controversy about NAPLAN has never really abated.

Last year's Gonski review signalled a need for change when it advocated for learning progressions and called for a way of assessing student progress that is more effective and timelier than NAPLAN.

This time last year, I wrote for the *Education Today* about learning progressions and I stand by my advice that learning progressions need to be empirically determined and that the current work on developing learning progressions needs to avoid mistakes made nationally and internationally with previous attempts to identify and articulate development in learning.

I would like to focus now on the

recommendation for a new online and on-demand student learning assessment tool based on the Australian Curriculum learning progressions. It's my understanding that ESA (Educational Services Australia) and ACARA (The Australian Curriculum and Reporting Authority) are working closely to develop this platform.

Copious amounts have been written about assessment in the last few decades, as we have grappled with trying to understand what assessment is, what we need it for and what we want it to do. We have invented new terminology to try articulate our thinking and phrases such as *formative and summative assessment, assessment for learning, assessment of learning, assessment as learning*, are now common parlance in education.





NAPLAN's death knell has not been sounded, but it will certainly change. Will standardised testing of the type we have known for the past 30 years, or so, go away? Will it be replaced by low-stakes classroom tests? Perhaps, for a bit. Maybe not. Education is the largest spender of the public purse. The need for teachers, for schools, for governments to be accountable is not likely to go away.

What if in the scurry to develop an on-line and on demand assessment tool, someone came up with a platform that enables teachers to easily create their own assessments, be it multiple-choice questions or rubrics and then to produce reports for school leaders and parents? Would that platform allow us to conceive of a new school assessment and accountability system?

No – no it wouldn't.

Let me explain why. I think we would all agree that learning is developmental, hence our desire to develop learning progressions. The conception of learning being developmental is readily evident in how we structure education. We don't start teaching algebra and work towards teaching students how to add two plus two. We don't start by teaching children to spell *Presbyterian* and go on to teaching them how to sound out *dog*. We start teaching elementary skills and then help students build upon these. Great thinkers such as Piaget and Vygotsky describe learning as being developmental.

Assessment's role is to tell us which students understand more, which students have more ability, which students

are more skilled, relative to the developmental continua. But if assessment is to do this, we need to first empirically determine the continua and then we need to carefully design assessments so that they provide insightful information of student performance in relation to those continua.

In the last few years, the saying *A year's growth, for a year's teaching* has taken hold in the educational literature and there appears to be a growing urgency for teachers to not only know where students are in their learning but to also know how much progress they are making.

From time to time, I lecture in educational measurement and I often open with "We need to measure student ability in the same way that we measure in the physical sciences. We need the educational equivalent of a thermometer to measure students' learning." Of course, everyone balks at that. The general consensus is that understanding student learning is far more complex than understanding temperature.

But here's an interesting paradox. Take a look at who worked on developing the thermometer and how long it took them. In 1593, Galileo Galilei developed a thermometer that could show changes in temperature. Ferdinand II, the Grand Duke of Tuscany, Daniel Fahrenheit and Anders Celsius all waded in, each grappling with ways to measure temperature. But it was not until 1867 that the medical thermometer was developed by Sir Thomas Albut, nearly 300 years after Galileo's breakthrough.

If measuring student ability is more

complex than measuring temperature, why do we think it can be so easy to generate our assessment instruments?

There's another irony. We rage against standardised tests and criticise the quality of the tests but we are very accepting of assessments created by teachers. I can speak with some knowledge of NAPLAN. All NAPLAN items are scrutinised by panels around the country and they are trialled extensively. Only a subset of items makes it into the final tests.

Please don't get me wrong – I am not being disparaging of teachers and our work shows teachers make great judgements about student performance, when you get the assessment process right. But think about the amount of time, investment and expertise that goes into creating standardised tests compared to the time that teachers have to create assessments. Teachers don't have time, not that kind of time, and they may not have the necessary level of knowledge, either.

The NAPLAN items are high quality assessment items. The issue is not so much their quality, but rather that only narrow aspects of the curriculum can be assessed through multiple-choice items. And I share many people's concerns that there is an over-reliance on the data collected through the standardised tests.

Let's be clear about what's right and what's wrong with our current standardised assessment. I think that there is a real risk that we will dismiss the standardised tests, and then devise systems that are on demand and easy to use, but which provide poor quality



assessment items. As the Grattan institute pointed out a few years ago, 'Poor quality tools will produce poor quality evidence, which is a weak foundation for targeted teaching.' Poorly devised tools are likely to divert student learning rather than promote it.

I firmly believe that teachers sharing banks of assessment items and rubrics will not take us in the right direction. We need to provide assessments that are of the same quality, or of an even higher quality, than we presently have in our assessment programs. But we also need to overcome any narrowing of the curriculum and any over-reliance on a single data set. Delivering items that are of the quality of NAPLAN items in a way that provides schools with flexibility is a great start and will likely help lessen the over-reliance on data, which is presently only collected once a year from students in Years 3, 5, 7 and 9.

But what are we to do about the narrowing of the curriculum?

There are inherent limits in what we can assess through multiple-choice items. We still need to find ways of assessing broader aspects of learning.

While there's promising work

happening around the world, it is surprising we haven't got further. Twenty years ago, Robinson recommended that the then UK Department for Education develop assessments appropriate for creative and cultural education. Clearly creativity in education, or the lack of it, concerns the millions of people who have watched Robinson's TED talks.

Why aren't we getting anywhere – in real terms, with finding ways of assessing creativity, critical thinking, problem solving, collaboration?

In his book *Creative Schools*, Robinson recounts a discussion he had with Andreas Schleicher, Director for Education at the OECD. According to Robinson, Schleicher said: "We always have to balance what is important to assess and what is feasible to assess."

Schleicher summed up the quandary, the quandary that is actually holding us back in educational assessment, if not education itself: 'Open-ended tasks are less reliable. You need human raters. You have the issue of inter-rater reliability.' Schleicher goes on to say, 'People don't like the open-ended tasks, because it's more expensive and it's a bit more

contestable, **but,** and this is an important but that Schleicher makes, 'on balance you get a lot more relevant information from open-ended tasks.'

This is the nub. This is the breakthrough that we are looking for. We need a way for teachers to reliably assess open-ended tasks that's relatively inexpensive. We need it because it will allow us to assess all the aspects of learning that can't be assessed through multiple-choice questions; the aspects of learning that relate to creativity and critical thinking. We need it because as Schleicher says, we will get a lot more relevant information about student learning.

The SMH headline spoke of, 'Cutting Edge' work. It is an exciting time as we re-envisage what a future assessment and accountability system may look like. Researchers, psychometricians and the education profession more generally need to do some hard work first up:

- We need to develop banks of high-quality assessment items
- We need to solve the issues associated with assessing extended performances
- If we want to measure and report on progress, we need assessments that provide a measurement scale.

Once we have done that, we can turn to the exciting advances in technology to make these assessments readily available to classroom teachers.

Like the views of Robinson and others, I have a strong belief that assessment in its current form is holding education back. I have an equally strong view that assessment is the key that will allow us to achieve much that we desire in education.

Dr Sandy Heldsinger co-founded Brightpath with Dr Stephen Humphry. Brightpath is an innovative approach to assessment and reporting, and is the result of over a decade of research at UWA to find a way of obtaining reliable teacher judgements of open-ended tasks.

Sandy co-ordinated the WA system-level assessments, has taught masters level courses in educational assessment for a number of years and has led the development of a wide range of resources, including reporting software, to support schools in using assessment to improve student performance. Sandy was acknowledged as WA's pre-eminent educational leader by the Australian Council of Educational Leaders in 2018.