



#21394491

What's the Point of Assessment?

The data generated by assessment doesn't necessarily improve student outcomes or teaching. Here's how a teaching team can plan a differentiated curriculum that hits a bullseye for both.

Ben Lawless

Aitken College, Lawless Learning and Matilda Education

I am impressed that the History Teachers' Association of Victoria has dedicated not just one article but an entire issue of this professional publication to the topic of assessment. Considering how assessment has been used in schools for decades, with confused purposes that contradict each other and outlandish increases in teacher workload that often amount to little improvement in student outcomes, it is no surprise that few teachers enjoy this topic. Indeed, I'm pleasantly surprised any time I speak on assessment and find that the room is not empty.

I'll open the issue with a few comments about terminology.

Formative vs Summative Assessment

You will often hear about the difference between formative assessment and summative assessment. The oft-quoted analogy is that formative assessment is when the chef tastes the soup, while summative assessment is when the customer tastes the soup. The point is that formative assessment can change the soup (i.e. student outcome) whereas summative assessment generally doesn't.

But this analogy is bogus. If all your customers complained that the soup was awful, might it not be changed? Second, assessments themselves are not formative or summative;

**In my opinion,
formative
assessment
is useful and
summative
assessment is
not.**

it is the way that we use the information that can be formative or summative.

Imagine if I set a Medieval History test in Week 1 of Term 1, and then leave out bits of the curriculum most students already know about. I am using the information from that test in a formative way—it influences my future teaching. If, however, I set a topic test at the very end of a unit, feed the results into my school's database and move to another topic, I am using the information summatively—it sums up how much a student knows about a topic.

In my opinion, formative assessment is useful and summative assessment is not. Formative use of assessment data can be used to improve your teaching. Don't teach students what they already know. Don't teach them things that are not developmentally appropriate for them. Formative use of assessment information is the first step to differentiating teaching based on ability.

Summative use of assessment information is for things like school reports and ATARs. Do school reports and ATARs improve learning? In one negative sense they might. While students might put in more effort to get a good report or ATAR, there are many examples of the ATAR system, for example, causing students anxiety at a very delicate time in their lives.¹

The purpose of assessment in education has been murky for a while now. Perhaps, in some bygone era, assessment was simply used to weed out the less academic types who would go on to a lifetime of fulfilling skilled or semi-skilled labour, leaving limited places in senior secondary school for students destined for the professions.

In other industries, assessment is used as information. How many cars is my factory producing each day? How can I increase that number? Assess the variables and try and improve in one of them. Assessment in these situations is used as information with which to make decisions to improve a process.

Unfortunately, for too long assessment has been used as judgement of students, teachers and schools—most of it negative. Assessment has a justifiably bad reputation in education because its use is seen as disparaging, and it is rarely used to improve processes.

Big Data vs Small Data

There is a lot of buzzing about 'big data' in the corporate world, and as managerialism (the idea that professional managers, not specific experts, should run things) has begun to take over the education system, the spectre of big data has fallen onto schools. However, big data is almost always irrelevant to the classroom teacher. My advice is not to worry about it. Even NAPLAN results cannot be used to infer student growth as the measurement error is larger than the expected growth.² What is important is the 'small data' that you generate yourself.

Assessment data should, in its best sense, only be generated when you have already thought of something you want to do with it. Unfortunately, the government requires schools to rate students on a minimum five-point scale at least twice per year.³ This is the ultimate kind of summative assessment data—it isn't really used to improve student outcomes or teaching. How could it? If a student gets an 'A' or '80 per cent' on an assessment, what information does it give teachers or learners? Get an 'A' next time? Try and get the remaining 20 per cent?

We have two kinds of assessment data—qualitative and quantitative. Put simply, quantitative data might be '80 per cent' while qualitative data might be that a student 'is able to multiply numbers up to 100'. For similar reasons to those already discussed, qualitative assessment information is much more useful than quantitative assessment information.

Why is there still so much quantitative data generated? It might be because it's easier to gather and is what has always been done in the past. It might be to flag whether an entire school is struggling. These are not good reasons to slow in-school assessment in this direction.

- 1 Tracey Bowden, 'Year 12 Exams—Are They Worth the Stress?' *ABC News*, October 11, 2017. <https://www.abc.net.au/news/2017-10-11/year-12-exams-are-they-worth-the-stress/9029260>; Deborah Corrigan, 'COVID-19 and Year 12 Anxiety: Should the ATAR Be Scrapped?' *Monash Lens*, April 2, 2020, <https://lens.monash.edu/@education/2020/04/02/1379964/covid-19-and-year-12-anxiety-should-the-atar-be-scrapped>; John Fischetti, Maxwell Smith and Raju Varanasi, 'We Know by Year 11 What Mark Students Will Get in Year 12. Do We Still Need a Stressful Exam?' *The Conversation*, July 27, 2020, <https://theconversation.com/we-know-by-year-11-what-mark-students-will-get-in-year-12-do-we-still-need-a-stressful-exam-140746>.
- 2 Margaret Wu and David Hornsby, 'Inappropriate Uses of NAPLAN Results,' *Practically Primary* 19:2 (August 2014): 16.
- 3 *Australian Education Act, 2013*, Australian Government, <https://www.legislation.gov.au/Details/C2013A00067>.

Norm-referencing vs Criterion Referencing

There are a couple of common ways to assess learners—against each other (norm-referencing) or against set standards (standards referencing or criterion referencing). By now you can figure out which one I prefer. Are we going to improve the learning of students by pitting them against each other, or by showing them what they are capable of doing?

What's the Purpose of Assessment?

What, then, is the point of assessing students? An often-heard trope is that it doesn't matter how many times you weigh a pig, it doesn't get any fatter. However, one would hope that at least some fattening of the pig is occurring between our students' trips to the assessment scales.

There's a couple of questions we need to ask ourselves first.

- Are we going to be able to interpret the results of our assessment meaningfully? If a student gets a 'B', what does that mean? How can I decode the information contained in the 'B'?
- Is the assessment going to improve student learning or improve how I teach that student? If the answer to both of those questions is 'no', what is the point?

Reporting on current achievement levels to parents and the government is obviously one common answer. However, this isn't a great use of student and teacher time. A related question is whether you already know the answer you're seeking. In many instances, students and teachers know in advance what the results of an assessment are going to be. Therefore, again, what is the point?

Differentiation

This brings me to what I would argue is the best use of assessment information: to diagnose a student's 'point of readiness' or Zone of Proximal Development.⁴ Once you know the student's 'Goldilocks zone'—not too hard, not too easy—teach them from that point. Some research

suggests <reference?> that up to 80 per cent of what comes across the average secondary school student's desk is either too easy or too hard.

There are some obvious criticisms of targeting instruction at the point of readiness, or 'differentiation by ability'. One major one that deserves our attention is that it's just too hard and/or time-consuming.⁵ I completely agree about this when teachers work by themselves. However, 'collective teacher efficacy' is the factor that has the highest effect size of them all.⁶ Here is a suggestion for how a teaching team of five could plan a differentiated curriculum, for both content and skills, for a ten-week term of history with three lessons per week. Let's assume:

- you're already using a textbook (which qualifies as the mid-level literacy document)
- you're teaching five historical skills
- you've mapped out five levels of sophistication for each historical skill
- this creates twenty-five quality criteria
- it takes about 30 minutes to write a targeted activity of the kind I've written⁷
- for content knowledge written at a lower literacy level you could either rewrite the textbook chapter (time-consuming) or find something online (quick and easy)
 - five minutes per piece of knowledge
- for content knowledge written at a higher literacy level you can find a more complex article about the topic from a free online encyclopaedia (e.g. Wikipedia provides great in-depth and expert-vetted materials) or another authoritative online source
 - five minutes per piece of knowledge.

Therefore you need:

- twenty targeted activities, not twenty-five—according to the rubric-writing guidelines I adhere to, the lowest level skill should be something everyone in your class can already do

Are we going to improve the learning of students by pitting them against each other, or by showing them what they are capable of doing?

4 Lev Vygotsky, *Thought and Language*, edited by Eugenia Hanfmann and Gertrude Vakar (Cambridge, MA: MIT Press, 1962).

5 Greg Ashman, *The Power of Explicit Teaching and Direct Instruction* (Great Britain: Corwin, 2020).

6 Hattie, *Visible Learning*.

7 Ben Lawless *Composite Rubric—History*, <https://lawlesslearningland.files.wordpress.com/2020/03/composite-rubric-history.pdf>

- thirty pieces of content knowledge written (or, more likely, sourced) at a lower literacy level
- thirty pieces of content knowledge written (or, more likely, sourced) at a higher literacy level.

| | | |
|---------------|--------|---|
| Task 1 | 60 min | Write two targeted activities |
| Task 2 | 10 min | Source three pieces of content knowledge at a lower literacy level |
| Task 3 | 10 min | Source three pieces of content knowledge at a higher literacy level |
| TOTAL | 80 min | |

8 How effect sizes are calculated is explained at <https://www.youtube.com/watch?v=IMGLzASbjYI>; a spreadsheet that will do the calculations can be downloaded at <https://evidenceforlearning.org.au/assets/Evidence-Exchange/Effect-Size-Calculator-Excel.xlsx>

Thus, for just eighty minutes per teacher, in two years **<please explain the maths!>** your team of five teachers would have developed a fully differentiated curriculum, with three levels of literacy and five levels of skills covered. If you increase the time that teachers spend on shared planning for differentiation by using an ‘assessment’ or

‘PD’ day, you could do an entire year’s worth of differentiated planning in a single day.

Pre-testing and Post-testing

Another assessment practice that is lacking in the History space is pre-testing and post-testing. Content pre-tests and post-tests can be useful to help us find out what students didn’t learn, and therefore allow us to do some reteaching if needed.

Perhaps even more useful for improving teaching effectiveness is content-agnostic skills pre-testing and post-testing. This is where you create a test that assesses student history skills (e.g. the big five from the Victorian Curriculum) but doesn’t require any content knowledge or use questions that have all the information needed to answer them. In this case you pre-test at the start of the History teaching part of the year, and post-test using the same test at the end of the History teaching part of the year. From this you can generate an ‘effect size’ that tells you how much impact your instruction has had.⁸ With this system you can find out which teachers are the most effective



MATILDA AD TO COME

and see whether anything they're doing is transferable to other teachers.

A really important message is that teaching is a skill, and some people are better at it than others. We can't hide from this fact. Just as we put our learners through a cycle of performance, judgement, feedback and improvement, we should be courageous and confident enough to expect it ourselves. Effect sizes are perhaps most powerful in a school context at comparing year-level courses. If your Year 7 results show that students developed a lot in their source analysis, but Year 8 students did not, that could be a good opportunity to ask questions about the two courses. What is it in the Year 7 course that helps student develop that skill?

The Final Word

There is more to say about assessment than can be discussed here. Let me just leave you with these opinions.

- 1. Formative use of assessment has a much bigger impact on

learning than a summative use of assessment information.

- 2. Assessment should be used as information to improve teaching or make other education-relevant decisions, not to judge students needlessly.
- 3. Teachers should be detectives and use 'small data' assessment to find things out that they need to know in order to improve a student's learning.
- 4. Assessment should be about what students can do, not how they compare with others. Perhaps there are instances in society when comparison is relevant, but these comparisons are demotivating and stressful in most schooling situations.
- 5. Use assessment to target instruction.
- 6. Differentiation is not possible without shared planning.
- 7. Pre-testing and post-testing of skills can be used to determine teacher and, more usefully, course effectiveness.

Ben Lawless' bank of targeted differentiated activities can be found at <https://lawlesslearning.com/free/compositerubrics/>

PROFESSIONAL READING ESPECIALLY FOR YOU

Classroom activities and in-depth history content written for Victoria's History teachers.

HTAV members can find back issues online.



htav.asn.au/agora


