

Unleashing the Power of Assessment: Placing Assessment Bank in the Hands of Teachers

Ian Anderson
Western Australian Primary Principals' Association

Sandy Heldsinger
The University of Western Australia

Stephen Humphry
The University of Western Australia

This paper recounts state and national accountability initiatives over the last thirty years and examines the impact of those initiatives on Western Australian schools. A 2010 survey of Western Australian educators found that one of the areas of greatest need was support for assessment. The paper explores research undertaken through joint research between Western Australian Primary Principals' Association and the University of Western Australia to develop an assessment process that uses teachers' assessments of their own students to meet their school accountability requirements.

Keywords: outcomes-based education, standardised testing, assessment for learning, teacher judgement

INTRODUCTION

Exciting and *assessment*: two words you don't often see together. But do have we an exciting assessment story to share with you?

A decade of close collaboration between the Western Australian Primary Principals' Association (WAPPA) and researchers from the University of Western Australian (UWA), has led to the development of a pioneering assessment process.

We want to share our work because the assessment process values teacher judgement. It rightfully places teacher assessment back within our school accountability system. And we are particularly excited about the data that show this innovative way of assessing is leading to improved student learning. Data collected over the past three years, show that students in high-usage Brightpath schools have progressed substantially more than other students, specifically three months greater progress over a two-year period on the Writing assessment in the National Assessment Program—Literacy and Numeracy.

Our story most probably begins back in the mid-90s when Western Australia (WA) followed a national and international trend to try to identify and articulate development in learning. Because of concerns for Australia's economic future in the late 80s, John Dawkins, the Federal Education Minister, had applied pressure on the Australian states to construct a common national curriculum. The states

obliged and developed profiles of learning. The profiles set out a hierarchy of learning outcomes and the intent was to assist teachers with identifying where students were in their learning and what they needed to learn next. As is often the case in our federation, each state and territory went on to develop their own versions of the profiles or learning progressions (Donnelly, 2007; Yates & Collins, 2008).

Before we continue with our story, which is a bit of a tale of woe, at least to begin with, we need to reassure you that WA schools are chock-a-block with knowledgeable, dedicated and professional leaders and teachers.

Outcomes-based Education

WA began work in earnest in the mid-90s to develop profiles of learning and it named its version of the profiles, the *Student Outcome Statements*. Later the name changed to *The Outcomes and Standards Framework* and somewhere along the line the term *outcomes-based education* took hold. This quickly became known by its acronym, *OBE*.

By 2004, 2005 things got pretty nasty in education in WA, so much so that the WA education minister was removed from her position. Mention *OBE* now and you can still see a nervous twitch ripple through the education community. With over 10 years distance, it's clearer now that the battle lines were drawn between those educationalists who valued a curriculum that was developmental in nature and those educationalists who were deeply frustrated by an assessment system that was cumbersome, time-consuming and that did not allow them to distinguish between students of differing ability.

The WA initiative was costly. Teachers had to be released from their classrooms *en masse* so that they could attempt to reach a common interpretation of the descriptions and ensure that their assessments against the levels were comparable.

Professor David Andrich, a WA academic and an international expert in educational measurement, was called in to review the crisis. Andrich found that the outcomes frameworks were too general and lacked the specificity required for developing teaching programs and for precise assessment (Andrich, 2006; Berlach & McNaught, 2007).

Australia is not the only example of failed learning progressions. The National Curriculum for England and Wales, which was introduced in 1988, consisted of levels or progressions. By the early 2000's a Commission called *The Commission for Assessment without Levels* was formed (Department for Education & Standards and Testing Agency, 2015).

The commission was damning of the National Curriculum. The levels and attainment targets, it concluded, had had a profoundly negative impact on teaching and learning. It found that the pressure generated by the use of levels in the UK accountability system had led to a curriculum driven by attainment targets, levels, sub-levels. The report stated that teachers had become focussed on getting students to the next level instead of ensuring their learning was secure and that they interpreted the levels differently so it was hard to tell what students really knew. It went on to say that parents and pupils did not really understand what it meant to be at a particular level but nonetheless used the levels to draw comparisons.

The commission's findings sound so similar to our experience in WA.

Standardised Testing

Australia followed another international trend in the 90s. A federal ministerial decree required all states and territories implement externally imposed standardised testing programs. The WA tests assessed 8, 10 and 12-year-old students' literacy and numeracy skills and a little later the program was extended to include students in the second year of secondary school. Like standardised testing programs around the world, students' reading and numeracy skills were assessed through multiple-choice questions. The state however committed to the more expensive, but valid option of assessing writing through students' extended written performances.

As part of the federal government initiative, a national project was established to equate the states' tests so that nationally comparable data about the number students meeting national benchmark standards could be reported. After several years, and limited success, the ministers agreed instead to implement a

national test. The National Assessment of Literacy and Numeracy, more commonly referred to be its abbreviation NAPLAN, was launched in 2008 (Rudd & Gillard, 2008).

As a side note, in December 2008, the Australian Ministers for Education published the Melbourne Declaration on Educational Goals for Young Australians in which they declared a commitment to improving education outcomes for all young Australians (MCEETYA, 2008). Of relevance to our story, the Ministers pledged that assessment of student progress would draw on a combination of the professional judgement of teachers and testing including national testing. Clearly, the ministers had already delivered on national testing before they declared their intent, but somewhere along the line their commitment to the professional judgement of teachers was lost.

NAPLAN – The Annual Behemoth

By 2019, the Australian newspapers were describing NAPLAN as a behemoth, and for good reason. Schools' performance on the tests is publicly available (Thompson, 2013). The data are colour-coded to show whether a school's mean performance is above or below the mean of schools with similar students. Green and dark green denote if the school's mean is above the similar schools' means and red and dark red to show if the school's mean is below the matched schools' means. Displaying the data in this way has led to the rather absurd goal of all schools trying to be above average and being displayed as a 'green' school.

Good work has been done across the states to help schools and teachers interpret their NAPLAN data and use the data in a formative and evaluative capacity. All this good work has been undone to some extent, because of the data being made public and the concomitant pressure on schools to perform well on the tests. In May 2019, our state newspaper ran the headline, '*Teaching to the test is making schools worse*', and many in education would agree (Carter, 2019).

Assessment for Learning

Whilst the NAPLAN juggernaut continues, Western Australian educators have been ardent followers of international academics whose research focus is formative assessment. Conference keynotes and workshops offered by such academics as Dylan Wiliam, Susan Brookhart, John Hattie, Andy Hargraves, Michael Fullan and Lynn Sharratt, have been very popular.

In 2010, WAPPA surveyed its members about the aspects of their work they found most challenging. Many reported that they needed help with assessment. The members reported a frustration about mixed messages and conflicting advice from policy makers and academics about assessment. They said that although the WA Department of Education required that all initiatives needed to be data driven, other than NAPLAN data, schools had limited access to data sets that they could use. The survey revealed a growing unease about uninformed interpretations of NAPLAN data and inappropriate claims by politicians and bureaucrats.

WAPPA and UWA's Vision – Dreaming the Impossible

Spurred on by the feedback from its members, WAPPA and UWA began to explore an alternative to externally imposed, standardized tests. Internationally there was a quest to find a solution that would enable schools to be accountable for outcomes in ways that were both valid and reliable. Policymakers in the US, for example, were considering formative assessment as a primary approach to educational reform (Heritage, 2010).

Our vision was to find a way of using teachers' assessments of their own students to meet their school accountability requirements. We were cognizant, however, that obtaining reliable data was an almost insurmountable stumbling block and that such an approach could add to the pressure on teachers (Humphry & Heldsinger, 2020).

So began our long research partnership which culminated in the development of a new assessment process in which teachers compare their students' performances to calibrated exemplars (Heldsinger & Humphry, 2010; Heldsinger & Humphry, 2013; Humphry & Heldsinger, 2019; Humphry, Heldsinger & Dawkins, 2017).

Fortunately for us this research coincided with advances in web-based software. We have developed software, which we have called Brightpath, so that teachers can more easily assess their students using the calibrated exemplars. The software also means that we are able to provide real time reporting, and importantly we can help teachers track and evaluate student growth in learning.

Take-up of Brightpath

Brightpath was launched in 2014 and in 2015 it was implemented as a significant component of the WA state testing program. After an extensive two-year trial with the Department for Education in South Australia, Brightpath is being used by the Department to support schools that have identified writing as focus of their school improvement plan. Schools from across Australia are using Brightpath and these schools range from high fee-paying schools such as Knox Grammar and Methodist Ladies College in Sydney and Ramininging Remote Community school in the Northern Territory.

Recently, the Federal Education Department asked Professor John Hattie and Professor Geoff Masters to recommend ways of supporting schools with formative assessment. Hattie and Masters identified Brightpath as one of the few tools aligned with well-constructed learning progressions and capable of providing information about the points students have reached in their learning and the growth they have made over time (Cawsey, Hattie & Masters, 2019).

Unleashing the Power of Assessment

As we mentioned earlier, data collected from WA schools over the past three years, show that students in high-usage Brightpath schools have progressed substantially more than other students.

School leaders describe the assessment process as in the moment professional learning. It appears the process of using Brightpath to assess, in and of itself, increases teachers' understanding of student learning and we want to investigate this further.

We believe that the following features of our work all contribute to informative assessment which in turn leads to more effective teaching:

- (i) teachers are well supported to make fine-grained assessments of their students and we think this helps them to teach writing more effectively,
- (ii) reports provide formative assessment information (success criteria, feedback to students, differentiated teaching etc.), and
- (iii) reports enable teachers and school leaders evaluate student growth in learning.

CONCLUSION

We began this brief recount of our work with a tale of woe. The WA experiment with articulating development in learning created a furore when teachers tried to use these descriptions to assess their students and monitor their progress. Although NAPLAN has now been in place for over a decade, it remains highly controversial and divisive. Teachers and school leaders wanted to follow the advice of leading academics but they didn't know how.

We weren't sure we could help but we wanted to, so we started to explore and research together and many WA educators have willingly helped us. We are excited with the progress we are making in Australian schools.

- We have demonstrated high levels of reliability where teachers assess their own students.
- We provide real-time assessment and immediate feedback based on empirical progressions and exemplars.
- Students in high-usage Brightpath schools have progressed substantially more than other students.

There is much that we still need to research and learn, but we believe that in one of the most remote cities in the world, we have uncovered a way of assessing that unleashes the power of assessment.

ACKNOWLEDGEMENTS

We would like to acknowledge and thank Stephen Breen and Katrina Bonetti for their central role in co-ordinating research in schools.

We would also like to acknowledge and thank the Association of Independent Schools of Western Australia for their contribution to our research.

REFERENCES

- Andrich, D. (2006). *A report to the Curriculum Council of Western Australia regarding assessment for tertiary selection*. Retrieved January 9, 2007, http://www.platowa.com/Official_word.html
- Berlach, R.G., & Mcnaught, K. (2007). Outcomes based education. Rethinking the provision of compulsory education in Western Australia. *Issues in Educational Research*, 17(1), 1-14.
- Carter, D. (2019, May 13). *Teaching to the test is making schools worse*. The Daily Telegraph. Retrieved from <https://thewest.com.au/opinion/teaching-to-the-test-is-making-schools-worse-ng-1b05c9e2ed69e52a7ce7c6d2bc8d0057>
- Cawsey, C., Hattie, J., & Masters, G. (2019). *Growth to Achievement: on-demand resources for teachers*. Australian Government, Department of Education, Skills and Employment. Retrieved from <https://docs.education.gov.au/documents/growth-achievement-demand-resources-teachers>
- Department for Education & Standards and Testing Agency. (2015). *Commission on assessment without levels: final report*. London, UK. Retrieved from <https://www.gov.uk/government/publications/commission-on-assessment-without-levels-final-report>
- Donnelly, K. (2007). Australia's adoption of outcomes-based education – a critique. *Issues in Educational Research*, 17(2), 183–206.
- Heldsinger, S.A., & Humphry, S.M. (2010). Using the method of pairwise comparison to obtain reliable teacher assessments. *Aust. Educ. Res.*, 37, 1–19. doi: 10.1007/BF03216919.
- Heldsinger, S.A., & Humphry, S.M. (2013). Using calibrated exemplars in the teacher-assessment of writing: an empirical study. *Educ. Res.*, 55, 219–235. doi: 10.1080/00131881.2013.825159.
- Heritage, M. (2010). *Formative assessment and next-generation assessment systems: are we losing an opportunity?* Paper prepared for the Council of Chief State School Officers, National Centre for Research on Evaluation, Standards and Student Testing (CRESST) Graduate School of Education and Information Studies, University of California, Los Angeles.
- Humphry, S.M., & Heldsinger, S.A. (2019). A two-stage method for classroom assessments of essay writing. *J. Educ. Meas.*, 56, 505–520. doi: 10.1111/jedm.12223.
- Humphry, S.M., & Heldsinger, S.A. (2020). *A two-stage method for obtaining reliable teacher assessments of writing*. *Frontiers in Education*. doi: <https://doi.org/10.3389/feduc.2020.00006>.
- Humphry, S.M., Heldsinger, S.A., & Dawkins, S. (2017). A two-stage assessment method for assessing oral language in early childhood. *Aust. J. Educ.*, 61, 124–140. doi: 10.1177/0004944117712777.
- MCEETYA. (2008). *Melbourne Declaration on Educational Goals for Young Australians*. Canberra, ACT. Retrieved from http://www.curriculum.edu.au/verve/_resources/National_Declaration_on_the_Educational_Goals_for_Young_Australians.pdf
- Rudd, K., & Gillard, J. (2008). *Quality Education: The case for an Education Revolution in our schools*. Canberra: Commonwealth of Australia.
- Thompson, G. (2013). NAPLAN, MySchool and Accountability: Teacher perceptions of the effects of testing. *The International Education Journal: Comparative Perspectives*, 12(2), 62–43.
- Yates, L., & Collins, C. (2008). *Australian Curriculum 1975 – 2005: what has happened to knowledge?* Paper presented at Australian Association for Research in Education Conference, Queensland University of Technology, December 2008. Retrieved from https://education.unimelb.edu.au/__data/assets/pdf_file/0009/974610/YAT081051_AARE2008.pdf