Assignment One

Research Contemporary Assessment Practices

by

Alan Hubbard

S326012

for

Dr Gary Fry

EST300 Assessment, Moderation and Reporting

Word Count: 1991

Assessing a student's learning comes in many forms, each with their own purpose. Whilst many forms of assessment can be used for a number of purposes, some are more suitable than others. Assessment at each level, classroom, state or district and national, is a contentious subject, that is constantly debated and under review to improve the process and pedagogies that support them.

Assessments can be defined in two main, albeit traditional, categories; formative and summative. Formative assessments are primarily used to improve an individual's performance and educate them. Generally, formative assessments can be used as assessment for learning. Summative assessments aim to assess what a student has learnt, the quality of their learning and is compared against predefined standards. Whilst this description means summative assessments are an assessment of learning, that does not mean they cannot be used as tools as and for learning (Dixson and Worrell, 2016). Summative assessments can be used as assessments for and as learning by allowing students to review the test material, their answers and correct answers, where possible, rather than awarding a grade with no elaborating comments.

The Northern Territory Curriculum Policy outlines the policy and guidelines for expectations towards assessment. In the Statement of Expectation – Assessment, this document outlines the schools are to assess, monitor and evaluate student's progress and achievement, which closely aligns toward the concept of assessing for, of and as learning. (NTBOS. n.d.) Assessment for learning is to allow teachers to identify where a student is in their learning journey, therefore allowing teachers to identify specific learning needs of students, for example diagnostic tests. Assessment of learning is to assess student achievement against the learning goals and curriculum standards, for example summative

assessment or examination. Assessment as learning engages students within the assessment process and allows them to reflect and monitor their own learning.

It further states to use processes to support teachers' judgement, whilst this is not explicitly explained, the implication is that this and other policies outline, clearly, what is expected of teachers and school leadership, and how to assess prepare students to take responsibility of their own learning journey.

Approaches by teachers, to classroom assessment varies depending on their career stage, experience with assessment and speciality. Further, many teachers are unable to conduct or participate in training on contemporary assessment models or techniques, therefore may be using ineffective techniques and models through a lack of professional development. Whilst, traditionally, teachers and pre-service teachers initially preferred summative and more formal formative assessment, such as self or peer assessment. After participating in teacher education programs, to improve assessment literacy, teachers tended to implement simple formative assessments, such as questioning and establishing success criteria and shared learning goals. This obvious shift in academic evidence away from traditional teaching models, is typically delayed through practice, due to the lack of professional development, training and education of teachers. (DeLuca, et al., 2018)

Darling-Hammond, et al., (2013) explains that based on demands by employers and the market on today and the future workforce, criteria to accurately assess these demands include assessments that encompass a wholistic approach to assessment for contemporary learners. These aspects are useful for reviewing current assessment techniques to better relate them contextually for students.

Assessment of high-order cognitive skills ensures that the skills learnt are transferable to other or new situations. High-order cognition is developed to be able to move beyond memorising information and to develop thought processes that give students the ability to identify problems or issues and develop methods to resolve or mitigate those issues, depending on the situation.

Critical abilities are key skills that students need to effectively operate in the modern workplace, such as communications, teamwork and self-reflection. Whilst subject matter is also key, without these abilities students will not be able to effectively convey their ideas, plan projects and reflect or self-assess their performance.

Due to the globalisation of the workplace, assessments should be assessed against international benchmarks so that their skills and achievements are easily recognisable internationally and against their peers and colleagues. Whilst some students will readily see the relevance of globalised standards, those that do not, should be guided toward that understanding.

Assessments should be designed so that the high-order cognitive skills and critical abilities are assessed rather than test-taking skills. The use of instructionally sensitive and educationally valuable assessments is often overlooked for traditional standardised tests that test how effectively a student or cohort has been taught rather than a student's capability. The assessment instruction should include valuable information that students can use to learn from and provide results that are useful for teachers to provide further instruction.

Additionally, quality of feedback from formal and informal assessment is important to enable the student to expand their current level of knowledge toward the targeted goals. Effective feedback needs to be immediate, specific and corrective, as necessary (Hafen, et al., 2015). As with some large-scale testing, due to a number of reasons, such as validation, confirmations and quality assurance, the results are often too broad or too late to be used for specific and targeted learning goals or plans.

Assessments must also be valid, reliable and fair to assess a wide range of learners and to assess a student's abilities across contexts. They should also support positive outcomes for student learning. This area critically relates to large scale assessments, such as National Assessment Program - Literacy and Numeracy (NAPLAN), where the test or assessment is designed for a specific purpose. Specifically, NAPLAN is designed for a specific purpose: "to provide an understanding of how individual students are performing at the time of the tests." (NAPLAN, n.d.)

NAPLAN data cannot provide accurate information about specific schools, cohorts or teacher performance, rather can be used to assess national education policy and framework (Lingard, et al., 2015) and to assess an individual's performance on particular skills at a point in time. (NAPLAN, n.d.) However, in a persistent need to achieve public accountability and confidence, highly publicised and politized test are implemented, regardless of the cost. Furthermore, whilst the intent of these assessments can, and should, provide informative data on the individuals and teaching practices that can be commended or improved, it has become clear that the information being used is producing negative outcomes. Such as the use of data to determine whether an individual class, and importantly their teacher, is performing to the prescribed standard. Firstly, this is not the intent of NAPLAN, therefore removes the validity

of the assessment for this purpose and can, create a negative perception of the teacher, their leadership, the school and the NAPLAN.

Gregson and Doidge (2018) raise the issue that whilst the governments from the United Kingdom, United States and South Korea place high emphasis on national standardised testing results, however teachers often and publicly dispute their efficacy in raising those results. Whilst the efficacy of standardised national testing is in dispute, without support by teachers and school leadership in its implementation, the national testing program is unlikely to succeed.

Reliability and precision of NAPLAN data is critical to the argument for or against the use of NAPLAN to inform policy and pedagogy to continual improvement. Wu (2015) elaborates on the reliability of NAPLAN reports, explaining how due to the poor test reliability, a student assessed to be at the mean level in Year 3 numeracy, could be as much as one standard deviation above or below that score. This means that an individual's numeracy competence could be as high as the top third of the assessed students and as low as the bottom third. This error is due to the breadth of questions, or lack thereof, that are available for a student to answer, being between 30 and 35 questions for each objectively scored assessment. Whilst this margin of error would be reduced by increasing the number of test items, it then may become a classroom or curriculum of testing rather than of learning. Furthermore, some or many students would be unable to persist through a longer testing regime.

Traditional models of testing, examination and assessment are still very prevalent in a number of teaching areas. These are typically used in the form of assessment of learning and

as a tool for placement and selection (Haertel, 2013-a). This continued use of exit or entrance examinations is one of the few pure forms of assessment of learning, where the student only receives a grade or percentage in the form of a result for that test or assessment, such as South Australian Certificate of Education (SACE) examiniations. However, the understanding that these tests measure a student's proficiency in academic work, rather than abilities in high-order and critical thinking, has changed significantly.

Whilst testing and the use of its data are very much contested, debated and disputed, their results have and can be used in a number of ways to improve the educational practices and pedagogies. Whilst it is not without flaws, through further improvements to this area of education, more reliable, valid and fair tests will become standard practice. While some test providers and users claim that only the intended effects of are their responsibility, all effects, intended or not, need to be addressed (Haertel, 2013-a). As with the extensive research and development that multiple choice tests has undergone, NAPLAN must also undergo wide ranging research and development, which will take a period of time before it is considered reliable by a majority of stakeholders, such as teachers, students parents and policy makers (Butler, 2018).

The Tennessee Comprehensive Assessment Program (TCAP) maths subtest was used to identify that not only maths comprehension, but specific subskills can be isolated as indicators for poorer future performance, at ages four to five. The evidence was clear that students with low maths comprehension pre-k, continued to score lower than those that scored higher. It was also shown that students that received retraining to target maths skills raised their maths comprehension to a similar level to those that initially received high results (Fyfe, et al., 2019). This use of a state-based standardised test is a key example of how

assessment of learning can be used to better enable educators to mitigate a lack of knowledge from previous education, to allow a greater chance of academic success during later learning stages.

The feedback loop or time taken for results and assessments to be received is an important part of the assessment process. If the time between assessment and receiving results is too protracted, the relevance of the feedback is reduced, as learners will have either forgotten the content of the assessment or will have moved past that point in their learning journey. This further reduces the assessment programs, in general, as some students will begin to be conditioned, that the assessments are only purposed as an assessment of learning with no relevance as a tool to enable their learning. In the case of assessments, such as NAPLAN, results are usually returned so long after the assessment date, that students cannot remember the content of the assessment or have learnt material which they did not know or fully comprehend since undertaking the NAPLAN bank of assessments. Similarly, SACE Year 12 or Stage 2 final examinations, do not provide students the information that shows where or what they answered incorrectly or insufficiently, and rather, only receive a grade or percentage, therefore cannot use the examination for learning.

Whilst large scale assessments cannot meet individual needs or specific requirements in all settings, it can provide useful data to aid teachers, schools, parents and other stakeholders or student education, to understand a groups' standards and understanding, therefore enable better policy decisions and pedagogical improvements. However, the data must be used for the purpose it was designed. State and classroom can provide better localised data, to enable smaller scale changes that may be affecting a specific group. Whilst assessment is unlikely to

ever leave the classroom, the assessment type and methods will change constantly, to meet

new research, policy and public opinion.

References

Butler, A. C. (2018). *Multiple-choice testing in education: Are the best practices for assessment also good for learning?.* Journal of Applied Research in Memory and Cognition, 7(3), 323-331.

Darling-Hammond, L., Herman, J., Pellegrino, J., Abedi, J., Aber, J. L., Baker, E., ... & Steele, C. M. (2013). *Criteria for high-quality assessment*. Stanford Center for Opportunity Policy in Education, 2, 171-192.

Dixson, D. D., & Worrell, F. C. (2016). *Formative and summative assessment in the classroom*. Theory into practice, 55(2), 153-159.

DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2018). *Teachers' approaches to classroom assessment: A large-scale survey*. Assessment in Education: Principles, Policy & Practice, 25(4), 355-375.

Fyfe, E. R., Rittle-Johnson, B., & Farran, D. C. (2019). *Predicting success on high-stakes math tests from preschool math measures among children from low-income homes*. Journal of Educational Psychology, 111(3), 402. http://dx.doi.org/10.1037/edu0000298

Gregson, R., & Doidge, N. (Eds.). (2018). *Connecting with Science Education*. Oxford University Press.

Haertel, E. (2013-a). *How is testing supposed to improve schooling?*. Measurement: interdisciplinary research and perspectives, 11(1-2), 1-18. https://doi.org/10.1080/15366367.2013.783752

Haertel, E. H. (2013-b). *Reliability and Validity of Inferences about Teachers Based on Student Scores.* William H. Angoff Memorial Lecture Series. Educational Testing Service.

Hafen, C. A., Hamre, B. K., Allen, J. P., Bell, C. A., Gitomer, D. H., & Pianta, R. C. (2015). *Teaching through interactions in secondary school classrooms: Revisiting the factor structure and practical application of the classroom assessment scoring system–secondary*. The Journal of early adolescence, 35(5-6), 651-680. https://doi.org/10.1177%2F0272431614537117

Lingard, B., Thompson, G., & Sellar, S. (Eds.). (2015). *National testing in schools: An Australian assessment*. Routled

National Assessment Program - Literacy and Numeracy (NAPLAN). (n.d.). *NAPLAN – general*. https://www.nap.edu.au/naplan/faqs/naplan-general

Northern Territory Board of Studies (NTBOS). (n.d.), *NTBOS Policy and Guidelines*. https://education.nt.gov.au/ntbos-policies-and-guidelines

Wu, M. (2015). *What national testing data can tell us.* In National Testing in Schools (pp. 18-29). Routledge.