

The effect of externally developed national testing in schools: Exploring two school sites

Katherine Swain

Donna Pendergast

Introduction

A shift towards a national assessment-driven approach to educational accountability occurred in 2000 with the Australian Government announcement that all education authorities should provide evidence of school standards through the reporting of literacy and numeracy performance against national benchmarks. The introduction of the National Assessment for Literacy and Numeracy (NAPLAN) in 2008 is one such measure, aligning as it does with a rise in high-stakes assessment programs around the world (Lingard, Martino & Rezai-Rashti, 2013). NAPLAN is intended to be a national accountability system that gauges student achievement and progress (Australian Curriculum Assessment and Reporting Authority (ACARA), 2010) and provide greater transparency and accountability for the performance of schools (Gillard, cited in Donnelly, 2010). NAPLAN involves standardised and norm-referenced testing of Australian students enrolled in Years 3, 5, 7 and 9 between ages eight and fifteen, in predominantly the middle years of schooling, annually in reading, writing, language conventions and numeracy. The tests are conducted in the school setting under strict administration principles and nationally agreed protocols to ensure integrity and consistency of test delivery (ACARA, 2011).

Subsequently the public website - www.myschool.edu.au - was launched to provide reports on NAPLAN outcomes for every Australian school, adding to the already high-stakes testing because of the links between NAPLAN results and government funding. The publication of school-level data increased the public profile of NAPLAN and consequentially changed the

way it was perceived. Researchers in the field of education such as Lobascher (2011), Polesel, Dulfer and Turnbull (2012) and Swain and Pendergast (2013, 2018) have noted that the introduction of NAPLAN signified a shift towards an assessment-driven approach to curriculum and an accountability-driven education system. It is in this space that our study set out to explore how Australia's national assessment affects curriculum, assessment and pedagogical practices in classrooms in two school sites. Specifically, we explored the impact of national assessment implementation on learning and teaching in seven classrooms in two school sites in Queensland. The study focussed on Year 7 students who were viewed as young adolescents, Year 5 students viewed as on the cusp of young adolescence and Year 3 students through the lens of those approaching adolescence. The school sites were selected because they employed different approaches to the implementation but shared location and socioeconomic status, with a similar demographic student profile.

Context

National assessment practice

In understanding the intent of national assessment, it is important to reflect on the purpose of assessment in general. Earl and Giles (2011) suggest the three main purposes of assessment are: assessment *for* learning, assessment *as* learning, and assessment *of* learning. Assessment *for* learning, includes all tasks undertaken by teachers and students to provide information which informs future learning. Assessment *as* learning involves students self-assessing to identify learning gaps as well as receiving and responding to feedback from their teachers. Assessment *of* learning, which occurs at the completion of a learning cycle or task and can be internal or external. NAPLAN is assessment *of* learning, according to this classification.

The approach that commonly underpins national testing is a psychometric approach to measuring knowledge or other attributes. Psychometric assessment uses standardised instruments designed to generate a measure of a perceived objective or attribute (Cumming, 2012). Norm-referenced approaches compare performance of a student to that of a group, and criterion-referenced standards can be used to profile an individual student's achievement in comparison to an expected standard. There are advantages and disadvantages to using psychometric testing (Cumming, 2012; Stiggins, 2005). Providing students, teachers and parents, with an indicator of how students compare to their peers, or to some criterion, is one of the advantages of psychometric testing. Another advantage is that it provides insight into what concepts need to be re-taught and reviewed (Brown, Irving & Keegan, 2007). Information gathered from psychometric testing is also used to make administrative decisions with regard to programs and other aspects of the system or to make decisions about the student (Gronlund & Waugh, 2008).

Despite widespread support from educational authorities of standardised tests, there have been numerous criticisms. Critics of psychometric testing argue that equitable assessment that strives to identify, as validly and reliably as possible, what students know and can do should offer students opportunity to demonstrate or perform what they know and can do (Rowntree, 1987). The Melbourne Declaration for Schooling (MCEETYA, 2008), advocates the need for appropriate educational practices in the middle years to ensure optimal learning outcomes for young Australians. Among the practices is a call for authentic and reflective assessment with high expectations, evidenced by higher-order thinking challenges (Swain & Pendergast, 2013) involving problem solving and reasoning (Kohn, 1999). Psychometric assessments do not engage these practices by challenging students to formulate their own

answers, nor do they further assess a student's knowledge base of a subject or directly assist student learning (Gronlund & Waugh, 2008; Stobard, 2008; Wyse & Torrance, 2009). Such tests, which are developed externally, especially national test programs which are developed and interpreted beyond the school site, decontextualises assessment from the classroom and halts learning when assessing (Broadfoot & Black, 2010; Gronlund & Waugh, 2008).

NAPLAN does not assess the full gamut of curriculum or developmental expectations of 21st century education. NAPLAN is not in its present form assessment *for* learning, which can support and promote learning in part because of the quality of interactive feedback (Black & Wiliam, 2009; Broadfoot & Black, 2010; Harlen, 2005). Formative assessment is part of the instructional process which informs both teachers and students about student understanding at a point when timely adjustments can be made (Ainsworth & Viegut, 2006; Black & Wiliam, 2009). This is an increasingly recognised issue, as formative assessment methods are known to be important to raising overall levels of student achievement. Indeed, as the Organisation for Economic Co-operation and Development notes, “[Q]uantitative and qualitative research on formative assessment has shown that it is perhaps one of the most important interventions for promoting high-performance ever studied” (OECD, 2008, p.2).

National assessment as high stakes

Accountability assessments are frequently described as high-stakes assessment. In Australia, NAPLAN is high stakes for schools and for state and territory education systems. Critics of national high-stakes assessment practices such as Cumming, Kimber and Wyatt-Smith (2011), Plank and Condliffe (2013) and Swain, Pendergast and Cumming (2018) warn of many unintended negative consequences of attaching high-stakes to assessment. These

include narrowing the curriculum, where teachers merely drill test content to achieve improved test results; where the curriculum lacks depth (teaching-to-the-test), where teachers coach students how to become successful test-takers (test wiseness) and where teachers focus learning and teaching on select groups of students to achieve maximum improvement, none of which involves high levels of academic rigour. Some preparation of students to participate in the types of tests is advised. Clearly, lack of prior experience could equally invalidate the student outcomes.

Broadfoot and Black (2010) propose that student achievement on high-stakes accountability assessments have become the legitimate currency for judging the quality of the education process and evaluating teacher performance. The pitfalls of such a practice to audit schools' performance and that of their pupils are many and therefore the use of such instruments for gauging the quality of teacher instruction is inappropriate (Broadfoot & Black, 2010).

Criticism of NAPLAN takes many forms. It is argued that NAPLAN testing need not lead to improved student outcomes. Among the many negative consequences of attaching high-stakes to assessment is that it can result in a narrowing of the curriculum where teachers merely drill content to focus only on those concepts and levels of cognitive skills required to achieve improved test results (Harlen, 2005; Shepard, 2003; Wyse & Torrance, 2009). Swain, Pendergast and Cumming (2018), warn that assessment systems lose much of their dependability and credibility when high stakes are attached to them. Indeed, in high-stakes testing regimes it is common for teachers to adopt surface rote teaching where regurgitation of mere facts is the outcome and the curriculum lacks depth and complex knowledge required for problem solving and decision-making (Harlen & Deakin-Crick, 2003) and explicit

scaffolding of learning experiences in which students participate (Hardy, 2013). Hardy (2013) in his research involving a school in north Queensland concluded that NAPLAN was impacting teacher practice. He found that some teachers engaged in explicit teaching about NAPLAN, including NAPLAN-like activities such as familiarity of test style and ensuring students were able to fill in the answer sheet correctly (Hardy, 2013). According to Hardy, such activities were ‘employed to further improve test results’ (2013, p. 75). A study conducted by Swain and Pendergast (2013) involving an analysis of NAPLAN reading tasks revealed that less than 10% of questions afforded the cognitive challenge of higher-order thinking, and hence the interests of middle years’ education and NAPLAN testing might offer some tension in our education system. Hardy (2013) warns that, ‘as a social act, the very process of counting necessarily influences, indeed ‘creates’, the world in which it is undertaken’ (p. 68).

A second and related concern is that, even if high stakes testing regimes improve performance in tests, they do not necessarily improve learning outcomes. That is, an increase in test scores is not indicative of an increase in student learning (William, 2008). Haladyna, Downing and Rodriguez (2002) and Swain, Pendergast and Cumming (2018) warn that test practising may lead to ‘test wiseness’ which will affect the consistency of the test results with repeated testing. ‘Test wiseness’ is therefore recognised as a threat to the validity of test score interpretation resulting in students achieving inflated results on skills where no mastery exists (Broadfoot & Black, 2010). A concern is that teachers may incorrectly read this ‘test wiseness’ as an indication of student learning, interpreting that their students have mastery of these skills and move on to the next level. Their students, however, may experience difficulty having not achieved depth of understanding of concepts and may be overlooked in terms of

potential intervention. A margin of error exists in all assessment (Stiggins, 2005), however, the results are often used as if the information is quite precise (William, 2008). This in turn misleads the teaching and learning process which highlights the consequences of misusing the data.

Hohensinn, Kubinger, Reif, Schleicher and Khorrarnadel (2011) argue the opposite of 'test wiseness' is 'test anxiety'. Broadfoot and Black (2010) note that, if anxiety affects test performance, it can be regarded as a source of invalidity which will distort the test scores. Hohensinn et al. (2011) similarly argue that students cannot perform to the best of their ability when they are upset or anxious. They further submit that a student's level of anxiety is dependent on the student's perception of cognitive demands. From this perspective, the use of the test scores may be problematic. Hence, the interpretation and use of test scores must be carefully considered when high stakes are attached.

Broadfoot and Black (2010) suggest that student's attitudes to learning and the strategies they use to further their own learning may be affected by the way assessment is conducted, therefore, if 'test anxiety' does interfere with optimum performance the anxiety might be reduced by making tests less 'test-like'. Whether 'test wiseness' or 'test anxiety' affects test results, students will only experience an increase in achievement through teaching (Black & William, 1998) and learning which involves higher-order thinking strategies and authentic and reflective assessment with high expectations as identified by the MYSA Position Paper (2008).

A further concern of the impact of high-stakes assessment is the narrowing of the focus of teaching. Black and Wiliam (1998) illustrate this effect using the analogy ‘it does not matter how much time the farmer spends measuring the pig; the pig will not get any fatter if the farmer does not feed it’. When teaching time is diverted to developing test skills and sitting tests, the time allocated to teaching is reduced. Critics of high-stakes assessment practices, such as Broadfoot and Black (2010) and Stobart (2008), support the concern that this culture of constantly measuring student performance reduces valuable teaching time and that the accountability movement that places inordinate value on test scores to ensure reaching a single benchmark will lead to the practice of ‘teaching to the test’ instead of teachers focussing on areas needing development and even neglecting the child. Plank and Condliffe emphasise this concern, noting that ‘Policies centered on high-stakes testing have, in many cases, achieved the goal of influencing day-to-day classroom activities’ (2013, p. 1153). When the nature of the assessment is high-stakes, the teacher pedagogy and learning experiences are subverted to mimic more closely the assessment with the result becoming more significant than the students taking the test. Such practices fail to use the cognitive skills of deductive and inductive reasoning, hypothesising, comparing, classifying and critiquing, all identified by Bloom (1956) as higher order thinking strategies. The exact same strategies recommended for inclusion in teaching pedagogies aimed at engaging middle school learners.

The New Taylorism

In this study, the lens of New Taylorism is applied to determine its harmony with the observations emanating from this study. New Taylorism evolved from Taylorism, which is the application of scientific management to increase efficiency. It was developed by mechanical engineer and author, Frederick Taylor, who in 1909 published a book that linked

factory productivity with work efficiencies (Littler, 1978). Taylor has been described as the ‘pioneer’ of scientific management and increasing productivity and his approach is still widely employed today, especially in business and management contexts (Turan, 2015). Taylorism is underpinned by four key principles, listed below, with language modernised for contemporary times by Caramela (2018):

- Break down assignments into subtasks
- Delegate responsibilities and train workers
- Monitor performance
- Allocate work between employers and managers.

Taylorism has been applied in a range of settings over the last century. The concept of the ‘New’ Taylorism is embedded in the original notion of Taylorism as the underlying approach that led to the US public education model. The original Taylorism model was designed to expand public education, making it available to the masses of students – accepted as a desirable goal. However, the addition of high stakes testing regimes makes this ‘new’ and one which Au warns serves the opposite function, that is to narrow education and place public education at risk while simultaneously “failing to prepare children for the intellectual rigors demanded within the globalised economy” (Au 2-11, p.40).

Au (2011) argues that high-stakes testing may lead to school leaders influencing classroom teachers to utilise standardised approaches such as focussing on test skills, scripted curriculum and narrow foci on what is being tested, as a response to the high stakes nature of national assessment. This is where teachers’ labour is “controlled vis-à-vis high-stakes testing and pre-packaged, corporate curricula aimed specifically at teaching to the tests” (p. 26). In this system school leaders lose autonomy in the way they conduct their usual pedagogy, impacting on the curriculum, pedagogy and assessment practices in classrooms.

The effects typically include curriculum narrowing and an increase in direct teaching, including drilling test content, practising testing protocols and test wiseness and do not include the comprehensive range of signature practices to engage young adolescents in relevant, meaningful and challenging learning (MYSA, 2008).

The New Taylorism lens is of particular interest in this study. Earlier research appears to reveal indications of this phenomenon occurring in the Australian context. For instance, a study of 8300 Australian teaching staff conducted by Dulfer et al. (2013) revealed a series of impacts that align with the notion of New Taylorism. For instance, teachers reported an impact on their teaching style and content choices, and that NAPLAN had led to a reduction in timetabling of other subjects in their schools; two-thirds reporting it led to less time to focus on other subjects. Teachers reported one of the many unintended consequences associated with NAPLAN is the narrowing of teaching strategies and of the curriculum. They reported that key learning areas such as: art, music and language were less likely to be addressed due to the increased testing focus. Furthermore, the study revealed that one-third of the teachers surveyed set more than seven practice tests prior to the testing period and 80 per cent reported test preparation added to an already overcrowded curriculum. Teachers reported a reduction in 'face-to-face' teaching time with one-half reporting that their pedagogy had changed and was more test-driven. Thirty nine per cent of the teachers reported they were teaching by rote and were administering weekly practice tests as a method of increasing test results. The findings indicate that NAPLAN has led to 'teaching to the test' whereby teachers were narrowing the curriculum in order to test children. Teachers also indicated that other unintended negative consequences included negative impacts on students' health and well-being, with 90 per cent of teachers reported students feeling stressed prior to

testing; staff morale, and school reputation and capacity to attract and retain students and staff (Dulfer et al., 2013). While such effects on students are beyond the scope of this study, classroom practices were firmly in our sights in order to ascertain if we also witnessed evidence of New Taylorism in action.

This study

School sites

The two primary school sites with students from Prep to Year 7 are located in south-east Queensland. They were selected because of their close proximity to each other and their socioeconomic and demographic similarity. However, their approach to NAPLAN differed at the school sites. In School A an intentional, highly defined program was delivered by teachers in the lead up and during the NAPLAN tests, including strategies developing test readiness, adjusted curriculum emphasis to favour literacy and numeracy, and adjusted teacher approaches to learning and teaching. In contrast, School B adopted a low-key approach to preparation for NAPLAN with no specific demands on teachers regarding preparation or attention leading up to or during the testing period.

Participants

The participants of the study were:

- the principal from each school,

7 teachers comprising 3 from School A and 4 from School B. (One teacher from School A was unable to participate in the data collection due to absence).

There were also 7 classroom observations; 3 in School A and 4 in School B. The observations were of the teacher participants' classrooms and sought to align the comments

about practice made by teachers in the focus group conversation with their actual practice in the classroom. The teachers and their year levels are summarised in Table 1.

Table 1: Study participants by school location

School A	School B
1 Principal	1 Principal
1 Year 3 teacher	2 Year 3 teachers
1 Year 5 teacher	1 Year 5 teacher
1 Year 7 teacher	1 Year 7 teacher

Data collection and analysis

The study design included three data collection stages. Stage 1 involved structured interview data collected from the Principals. The purpose of this was to gain an understanding of the school culture with respect to the implementation of NAPLAN. Stage 2 involved focus group interviews with seven of the eight classroom teachers. Stage 3 involved observation data collected from classrooms at each school site. The purpose of this stage was to investigate the alignment between teacher comments and actual classroom practice.

More than 70 pages of transcripts from Stage 1 and Stage 2 were analysed using thematic content analysis. “To make sense of the text” (Creswell, 2007, p. 244) the analysis and interpretation of transcriptions of students’, teachers’ and parents’ focus group interviews, formal interviews conducted with the school managers, and of observation checklists and field notes, follow Creswell’s six stage model for analysing and interpreting qualitative data.

1. Preparing and organising data

2. Exploring and coding
3. Describing findings and forming themes
4. Representing and reporting findings
5. Interpreting the meaning of findings
6. Validating the accuracy of findings

This paper will focus on Stage 2, the teachers' viewpoint and Stage 3 classroom observations.

Stage 2 - Teacher semi-structured focus group interviews

The teacher focus group interviews involved one focus group for each school. Teachers were encouraged to share experiences of curriculum, pedagogy and assessment practices in their classrooms leading up to, during and post the implementation of the testing process. Prompt questions were used to ensure the flow of conversation. Questions were designed around NAPLAN practices and procedures and fell into three categories: students; teachers; and pedagogy. Table 3 outlines the prompts used to guide the direction of conversations and the related category of each.

Table 3 Teacher focus group questions/prompts

Prompts/Questions	Category
Does NAPLAN affect your students? Explain	Students
Do you notice any change in student/teacher relationships or student behaviour during NAPLAN preparation period? Explain	Students
Do you think NAPLAN affects student attendance?	Students
What was your classroom like before the introduction of NAPLAN?	Pedagogy
What is your ideal teaching situation?	Pedagogy
Has preparation for NAPLAN affected student time in relation to: Excursions, Classroom enrichment activities, Student performances, Parent contact etc?	Pedagogy
Do you feel pressured by the parents to increase your class's NAPLAN results?	Teacher
Have you ever felt like transferring to a year level that does not have NAPLAN?	Teacher
How much time would you spend per year on NAPLAN practice?	Pedagogy
Do you think that practising for NAPLAN is the only way to achieve improved test results? Explain	Pedagogy
When you are preparing for NAPLAN do you target specific groups of students? Explain	Pedagogy
During the NAPLAN preparation period how much time would you spend on: individual seat work, whole group instruction, basic skills, concept development using hands on activities, critical thinking?	Pedagogy
Do you think that NAPLAN results impact you personally? Explain	Teacher
Do you feel that NAPLAN results will be used to award teachers and administrators financial bonuses?	Teacher
Does NAPLAN preparation only occur in those year levels involved in the testing?	Pedagogy
Do you think there are better ways of assessing student abilities or is this the best way to achieve valid results? Explain	Pedagogy

To support the data collected from the teachers, and to assess the level of alignment between participants' accounts of the learning and teaching environment, class observations were undertaken during the test preparation period.

Stage 3 - Classroom Observations

Classroom observations took place in the test preparation period prior to testing in May. Data were gathered using a criteria checklist. Seven separate visits spanning 45–60 minutes were conducted in the Years 3, 5 and 7 classrooms. The structure for the observation schedule involved four points of focus: the physical arrangement; classroom atmosphere; student and teacher activity; and lesson content, context and pedagogy.

Findings

The findings are presented separately for the two stages.

Stage 2 - Teacher semi-structured focus groups

Semi-structured focus group interviews were conducted with one group of teachers at each school. Questions/prompts were related to three categories, students, teachers, and pedagogy. Group discussion commenced in relation to implementation of NAPLAN and any perceived consequential impact on student/teacher relationships, student behaviour and student attendance. Having focused on possible effects of NAPLAN on students, the focus of topic turned to include possible effects on teachers personally and professionally. Teachers were also asked to compare learning and teaching at their schools prior to, and post, the implementation and discussed their perceptions of, and reactions to, a perceived pedagogical shift.

School A Teachers

Three teachers at School A attended the focus group interview, the fourth teacher was forced to withdraw due to illness. The teachers reported that preparation for NAPLAN commenced at School A in the second week of the school year with two-thirds of every day devoted to practising for the testing period in May.

The Year 7 teacher's perception of NAPLAN testing was somewhat dissimilar to that of the Year 3 and Year 5 teachers. The Year 7 teacher disputed that NAPLAN and preparation for NAPLAN affected her students negatively. She reported 'teaching to the test' and supported changes to classroom structure and teacher pedagogy to accommodate NAPLAN. In contrast, the Year 5 teacher, and the Year 3 teacher both perceived national testing and preparation as impacting students negatively. They suggested that it created unnecessary stress short-term and consumed valuable teaching time which resulted in long-term side-effects. They added that practising for the tests involved introduction of too many concepts over too short a time which reduced opportunities to appropriately scaffold student learning for development of deep understanding. This focus on test preparation caused a shift in their classrooms from student centered pedagogy conducive to best practice for middle school learners involving cooperative learning and collaborative teaching (MYSA, 2008; Swain & Pendergast, 2013) to a more teacher focussed approach.

School A's teachers did not support the inclusion of Year 3 students in NAPLAN on the grounds that they were too young to be included, although conducting practice tests did 'ease their suffering'. The Year 5 teacher reported, '[I]n Grade 3 they still need to learn those

logical steps. I do feel that it is going to affect their later learning because now they are not learning'. The Year 7 teacher refuted any adverse influences on Year 7 students and claimed that the benefits over-rode any negative outcomes. However, she later stated that the national testing can cause Year 7 students stress.

Concerns were raised by School A's teachers regarding restricting students to a limited duration to complete tests and that these time limits created unnecessary stress. Furthermore, the Year 7 teacher questioned the integrity of including trick questions in an attempt to confuse students. School A's teachers maintained that the tests were set above their students, the Year 5 teacher explained, '[I]t's way too high for them. It's crazy. There is a lot in the Reading test for them to read for the time they have'. The Year 3 teacher added, 'It's not aimed at our year level because there's a lot of inferential reading, there is no way you can do it in time'. School A's teachers suggested that trick questions were embedded in the tests and that the added pressure involved in teaching students how to identify a trick question consumed time better dedicated to more important tasks.

Continual test preparation was reported to impact School A's student/teacher relationships; however, teachers presented differing perceptions of resulting repercussions. The Year 7 teacher viewed changes in her classroom as positive for the purpose of easier behaviour management. She contended that test preparation positioned her students in test-like situations where interaction was disallowed, thus ensuring simplicity when detecting inappropriate student behaviour. She explained that the process of testing, analysing, addressing weaknesses and re-testing using past tests reassured students that their teachers would provide them with appropriate skills, knowledge and understanding in order to achieve

success. According to the Year 7 teacher this routine assisted in development of student/teacher relationships based on trust.

In contrast, the Years 3 and 5 teachers suggested that constant practising for NAPLAN resulted in insufficient time available to provide experiences which assist in development of a healthy school environment with an emphasis on strong teacher–student relationships (MYSA, 2008). The Year 3 teacher explained, ‘I just think they still need that TLC (Tender Loving Care) and they don't get that I don't have time...’ The Year 5 teacher added, ‘[F]irst term you should be focusing on behaviour and tuning the children into their learning, but it is really hard to engage the kids when from the start of the year it is testing’.

The teachers at School A admitted to feeling pressured by NAPLAN and all that it entails. Expectations placed on the principal by higher authorities were shared with teachers, who then passed them on to students. Year 3’s teacher explained that she gets anxious about how her students suffer and added that some of her children were crying. She explained, ‘[I]t doesn’t matter how you go about starting the test and preparing them for it they still feel the pressure’. The Year 7 teacher suggested a further stress related to externally developed tests involved teacher effectiveness being gauged by students’ results. If their students failed to achieve improved test results, they feared punishment by the principal by means of transfer to a Year level not involved in the testing. This meant that those teachers with a love of teaching middle years students were forced to teach in other year levels where their skills were not utilised. The Year 7 teacher explained that students’ test results were, ‘Pulled apart... scrutinised... year level by year level... teacher by teacher. The principal does graphs, and he puts them up in staff meetings and we all know who's who’. Teachers at School A claimed

that NAPLAN results were not indicative of best practice teaching designed to engage the young adolescent as there were other areas of curriculum equally as important as those addressed in the tests.

According to School A teachers, curriculum and pedagogy had significantly transformed since NAPLAN had been implemented. They outlined that learning and teaching prior to NAPLAN included whole term integrated unit plans which encompassed curriculum expectations across seven key learning areas. Highly scaffolded concepts were introduced through topics generated from the students' real world. There was time for explicit teaching of basic skills, school camps with curriculum connections, art and drama experiences, and learning through discovery and problem solving which included higher order thinking skills. All of these aspects had been eliminated and replaced with NAPLAN preparation involving 'teaching to the test' and 'test wiseness' thus narrowing curriculum.

The teachers revealed genuine concern for their students' future. The Year 5 teacher explained that, 'rather than doing one concept in depth in one day she probably covered five or 10 concepts.' Each participant identified the absence of important pedagogy such as hands-on manipulative material for students still functioning at the concrete operational stage. They added that the frequency of student group work had reduced and reported that most instruction was whole class and involved students sitting at their desks the majority of the time and not involving the integrated and disciplinary curricula that are negotiated, relevant and challenging suggested by MYSA (2013). The Year 3 teacher explained, 'We don't sit on the carpet and share; tell stories and run language programs.' School A's teachers suggested that curriculum, assessment and pedagogy conducive to best classroom practices included

appropriate time available to cover concepts from foundation through to deep understanding and was not that which was associated with NAPLAN structure.

School A's teachers suggested that student data collected from administering multiple practice tests were used during parent/teacher interviews. They surmised that multiple collections of results constituted a more reliable approach to obtaining data than the 'point-in-time' NAPLAN tests which produced invalid data because students were able to guess or students may be experiencing other issues which may affect their ability to achieve, which would skew results. Skewing of results raised concerns as they believed that NAPLAN results impacted the distribution of funding received from governments for intervention to address student needs. The Year 5 teacher described how skewed results can impact classroom practice. 'I am having issues with a student at the moment that is not getting certain support because she did okay on the last tests, but I think it was guessing'. Teachers suggested further point-in-time testing issues which may impact data validity included bright students having a bad day, students feeling sick and those suffering test-anxiety.

Another concern raised was with regard to lower ability students, who teachers described as disadvantaged. They explained that the students who do not initially understand a new concept should be exposed to other ways to learn that concept, but this is not happening as there is not the time available. School A's teachers explained that their principal expected improved student NAPLAN results and the easiest way to improve the classroom average is to target students who do not need to travel far to move from below to above the average line. At this crucial period of adolescence when cognitive processes are undergoing refinement (Pendergast, 2010) the- teachers at School A considered spending time with low achieving

students during test preparation to be a waste of time. Instead they reported their involvement in selecting students positioned close to, but just below state and/or national averages and targeting them with focused teaching in an attempt to elevate students' positions thus increasing class averages. The Year 3 teacher explained:

I gave my teacher aide the lows and I solely targeted the middle to high group and I did that every day. If any are going to get better results it will be those students...I know this is terrible, the really low ones who cannot read instructions to start with, there is no point me working with students that cannot read.

The care these teachers directed towards their students was obvious, but they described issues which caused them anxiety as they complied with changes to curriculum, pedagogy and assessment practices, which they identified as non-conducive to best classroom practice.

School B teachers

Four teachers at School B attended the focus group interview. NAPLAN implementation was described by these teachers as having had minimal impact on teaching and learning in their classrooms and there was no pressure from the Principal to make changes. They reported spending less than an hour a year rehearsing for NAPLAN. As traditional paper and pencil testing was not common practice at School B, teachers claimed that their students of all ages would be uncomfortable in a formal testing situation. The unfamiliar structure of NAPLAN and extensive testing period were identified by teachers as their primary concerns.

School B's teachers explained that test material was integrated within existing curriculum, thus camouflaging any changes to lesson structure or content. The Year 5 teacher explained,

I linked it in with the main lesson...they were exposed to it and I made it real for them so they felt confident with it...I got one test like one practice test and we cut it into strips and I made it into a game, so it was fun which is good for relationship building.

School B's teachers reported an absence of pressure in relation to NAPLAN and reasoned that it reflected the low level of importance NAPLAN was afforded at School B. It would appear that very little change had occurred in relation to test implementation except during the actual testing. Group, individual or whole class activities where students were offered opportunities to learn from each other and discover solutions to problems using manipulative materials were commonplace in these classrooms. The Year 5 teacher explained, '[W]e are about the holistic child, not just about one test'. According to these teachers, teaching and learning which existed in classrooms at School B were conducive to effective learning and involved authentic instruction (Lingard, 2003) which involves learning and teaching that excites students in twenty-first century classrooms by engaging them with work of intellectual quality (Horan, 2010; Lingard, 2003) where they are able to identify the links to their world. They suggested that placing students at the centre of curriculum presented opportunities for success and encouraged students to develop a love of learning. They added that assessment practices at the school were ongoing and in the main formative in nature which resulted in valid assessment of student achievement.

Teachers from School B agreed that NAPLAN did affect their students negatively by way of test anxiety and stress which could impact their demonstrations of achievement. They reported that some students were 'a bit stressed' and others, 'a bit scared'. One teacher described a parent pushing her daughter into the classroom to participate in the testing. The

Year 7 teacher questioned how children who are stressed and scared could concentrate on the testing.

A further issue raised by the four teachers at School B was the varying levels of test preparation conducted in schools and therefore was the testing conducted on a level playing field? They were aware that some teachers at other schools were ‘teaching to the test’ and teaching to specific groups of students so as to achieve improved results. These teachers stated that they did not engage in such practices.

School B’s teachers speculated that NAPLAN implementation may lead to the introduction of performance-based pay in the government sector which they feared would negatively impact the teaching profession including far reaching side-effects such as teachers cheating, pressuring students to perform, knowledge becoming power, and massive public sector resignations. School B’s teachers indicated that the way forward from this point was to continue with current structure, curriculum, pedagogy and assessment practised at School B because NAPLAN was, ‘Not the best way. In fact, it’s possibly the worst way’.

Stage 3 - Classroom observations

Classroom observations were conducted early in the year and prior to NAPLAN testing. The purpose was to observe the alignment or otherwise of what was reported by the teachers. Important indicators of the key components of middle schooling principles and practices provided a lens through which to view the middle years classrooms involved in the study. Focussing observations through the signifying practices of middle years criteria allowed the observer to determine whether or not these middle years classrooms were democratic

classrooms where knowledge was socially constructed, where the students had ownership of the classroom and were the centre of the curriculum, and where negotiation and integrated curriculum were common practices (Swain, 2015).

School A Classrooms

Classroom observations conducted at School A involved visits to Year 3, 5 and 7 classrooms.

Furniture placement varied in classrooms observed at School A. Noted in all rooms was the lack of useable floor space while electronic whiteboard positioning dictated a definite front to each room. Wall displays in all participating classrooms included Literacy and Numeracy posters and very few student work samples. Students' practice test results written next to students' names were displayed on the Year 7 classroom wall.

The whole class lesson observed in the Year 3 classroom identified curriculum intent as numeracy found on the national tests, however, teaching numeracy skills or concepts was not observed. Each practice test question was dissected, focussing on identifying clues in the question format which would assist students to discover correct answers and was not about solving numeracy problems. Instructions included such statements as, 'If there are two lines, then they expect two answers.' This further clarified lesson intent as 'test wiseness'.

Observation in the Year 5 class involved a Literacy test. Lesson content included strategies which aided students' attempts to identify the distracters supplied in NAPLAN multiple choice options for literacy. These included explanations such as, 'If a multiple-choice question ends in the word 'an' the answer must begin with a vowel. If you have no idea what the answer is, choose the longest one'. Other concepts covered during the lesson included: capitalisation; colouring small circles correctly; and 'the amount of spaces provided for the

answer determines the length of the answer required'. Each question involved a different concept, no links were identified between concepts and there was no opportunity for students to consolidate each concept before moving on. Several students faced away from the teacher including a boy sitting alone at the end of the teacher's desk who received no teacher acknowledgement the entire lesson. Observed in this classroom was 'teaching to the test' regardless of student engagement or disengagement.

The only classroom observation conducted in the afternoon session at School A was in the Year 7 classroom. Whole class instructions were given to students before they commenced working collaboratively identifying animal endangerment and its relationship with humans. Students reported that the previous two sessions had involved practice testing. Evidence of this was observed on entering the room to witness the teacher calling out students' names and students responding with their practice test results. As each result was supplied the teacher repeated it loudly to the whole class.

During the observation period no direct or indirect reference to NAPLAN was made in the Year 7 classroom. However, during the 45-minute observation conducted in Year 5 the word 'test' was recorded 17 times. Students in Years 3 and 7 were observed receiving verbal praise. In Year 5 verbal praise was forthcoming for students answering practice test questions correctly, however, those students who answered incorrectly were subjected to sarcasm and 'put-downs'. Student movement in these classrooms was limited.

Students were supplied with resources required to complete classroom tasks, but few other resources were observed in these classrooms. This is not to say that there were no other

resources available to students attending School A, but these were the only resources observed in classroom settings. Year 3 students also demonstrated disengagement through talking and daydreaming, only three students produced any work. Children in Year 5 displayed negative responses towards the teacher and task. Five students engaged in the learning experience. The lesson continued regardless of varying levels of student engagement. In the Year 7 classroom students were observed working collaboratively. Their teacher continually prompted students to extend their thinking. This classroom was unique in that it was the only classroom where students were observed working collaboratively with other students and their teacher.

These classrooms were not learner centred. The learning environment portrayed little collaborative activity where students could develop quality relationships. Pedagogical practices in these classrooms did not provide opportunities for students to identify with a sense of self, nor were the activities rigorous, practical real-life and relevant.

School B Classrooms

Observations were conducted at School B in two Year 3 classrooms, 3A and 3B, one Year 5 classroom, and one Year 7 classroom. Classrooms at School B included double teaching spaces which allowed room for furniture arrangement while providing floor space for individuals, small groups or whole class floor activities.

All classrooms at School B contained student desks and shelving. Observed desk arrangements included rows and group format. Students were not allocated specific places to sit nor was there designated ownership of desks. Students were recorded moving freely around rooms addressing tasks through their selected mode of learning. A definite front of room only existed in the Year 5 classroom, where the desks positioned in rows faced the

blackboard at the designated front of the room. Demonstrations of student work constituted the majority of items displayed on walls of all classrooms observed at School B.

On entry to the classrooms at School B it was difficult to locate teachers as they constantly engaged with students and tasks. Students in the Year 7 classroom and 3B were observed negotiating tasks and learning environments confidently with each other and their teachers. Children in 3A engaged in reading while their teacher provided individual consultation when required. Lesson content in the Year 7 classroom involved research, design and problem solving as students investigated famous inventors and inventions through history. The Year 7 teacher consulted with each student and appeared to offer encouragement and ideas and prompted their research. Students belonging to 3B, positioned in a circle on the floor, revisited their social skills, their teacher, also a member of the circle.

The format of mathematics observed in the Year 5 classroom resembled that included in NAPLAN, as the mathematical tasks and questions were presented in multiple choice format. Mathematical problems were presented and strategies for solving problems were discussed. Students were asked to select and justify their preferred responses from a selection of four multiple choice answers. Manipulative materials were available for those students who functioned at the concrete operational stage. Students were also given the opportunity to consolidate concepts introduced in each mathematical problem before moving on to a new concept. Noted was an absence of direct or indirect reference to the tests. Consistent in all classrooms observed at School B was encouragement and verbal praise directed towards students.

Available resources observed at School B included books, games and concrete manipulative materials positioned in labelled boxes situated around the perimeter of each classroom. Noted was the absence of computers in classroom settings. Students were instructed to relocate to the library if they required the use of information technology. Observations of 45 to 60 minutes are a snapshot of the classroom environment and do not offer an holistic view overall. During these observations the characteristics of best practice Middle Schooling (MYSA, 2008) were evident, and all students appeared to be engaged and motivated in the teaching and learning environment.

Discussion

The findings from Stage 2 focus group interviews with 7 teachers and Stage 3 seven classroom observations, provide a unique window to explore the effects on curriculum, assessment and pedagogical practices in NAPLAN classrooms in two school sites. Also, this stage of the data collection allowed for identification of any existing tension between the classroom practices at the two schools and the signifying practices of middle schooling (MYSA, 2008). School A allocated considerable class time to test readiness and emphasised the importance of NAPLAN while School B engaged in minimal preparation and did not highlight NAPLAN as an important aspect of the school experience.

The findings from the two stages of this study are now discussed via key themes that emerged from thematic content analysis of transcripts from Stage 2. The analysis of the transcribed data followed Creswell's Six Stage Coding Process (2007) which allows clear identification of major themes through constantly revisiting the data, revealing five main themes:

- hidden agendas and top-down pressure;
- NAPLAN as diagnostic or holistic;
- curriculum, assessment and pedagogical shift to a NAPLAN focus;
- NAPLAN practices a positive or negative;
- NAPLAN equity and validity.

Hidden agendas and top-down pressure

School A's teachers suggested that NAPLAN was a tool for judging teachers and schools and not for the purpose of increasing student achievement. They suggested that their perception of a hidden agenda extended beyond politicians and that their principal judged his teachers as successful or incompetent according to their students' NAPLAN results. Furthermore, they indicated failure to demonstrate student improvement led to punishment by being moved to a non-testing year level the following year. Teachers at School B concurred with this view, describing the assessment program as not being about children, but about 'political point-scoring' and money. Teachers at School B identified performance-based pay for teachers and management in government schools as the hidden political agenda. They suggested that parents rated schools according to the school's NAPLAN results and governments would use similar methods to determine salaries for government employees. This aligns with Au's (2011) conceptions of New Taylorism in terms of teachers efficiency and productivity in the form of student success.

Data indicated a commonality between the two research sites in that study participants perceived the existence of hidden agendas in relation to the development and implementation of NAPLAN and how the data it produced might be used. School A's teachers recognised that

their principal was pressured by superiors and this pressure was passed on to teachers, who in turn passed it on to students. Teachers at School B did not report experiencing top down pressure in regard to NAPLAN, as it was not prioritised at School B. Evidence of top-down pressure was not observed at School B which supported the teachers' accounts of occurrences at School B.

NAPLAN as diagnostic or holistic

One teacher at School A reported that testing and pre-testing had proven to be useful as a diagnostic tool for the identification of gaps in education of students. In this way constant exposure to problem solving strategies involved in practice tests had resulted in increased levels of student achievement. The remaining two teachers stated the only benefit students gained from constant testing and re-testing was related to familiarity of test format, processes and procedures.

The teachers at School B did not place value on NAPLAN as a diagnostic tool and did not choose to implement a constant test and re-test preparation focus as a tool for diagnostic assessment of student learning needs. They reported that teachers' judgments were a more reliable gauge of student achievement. School B's teachers added that the time period between testing and results delivery made the assessment ineffective as a diagnostic tool.

The teachers at School A did not regard NAPLAN to be a valuable holistic assessment tool arguing that literacy and numeracy did not alone provide all skills and knowledge required for a complete education. They agreed that students' results for achievement in the wealth of programs which ran in classrooms after NAPLAN was completed in May were equally as

important and were indicative of the value of education at School A. School B preferred overall assessment practices conducted by classroom teachers to gauge distance travelled by their students.

Curriculum, assessment and pedagogical shift to a national assessment focus

In order to address the perceived accountability agenda School A endorsed drastically altered curriculum, pedagogy and assessment practices in Years 3, 5 and 7 for four months leading up to NAPLAN and to a lesser degree of intensity, Years 2, 4 and 6 for six months in preparation for NAPLAN the following year. Teachers of Years 3, 5 and 7 at School A were instructed to replace their traditional curriculum, pedagogy and assessment practices with a rigorous program of constant testing and retesting. Tests from previous years were used to identify student weaknesses, teachers then taught to the weaknesses before implementing further testing to evaluate student progress and to identify need for intervention. These practices align with Au's (2011) explanation of New Taylorism that high stakes assessment directly impacts classroom practices, and not for the best.

Teachers at School A stated that the expectation was that they would comply with the assessment teaching focus and prescribed assessment tasks as they were not negotiable at School A. They reported that preparation for NAPLAN commenced in week one or two of each school year and curriculum and pedagogy focused on assessment which left no time for 'getting to know you' activities or for developing social skills, and classroom rules and expectations. Teachers reported that two-thirds of each day was devoted to completing practice tests, marking tests and analysing results which involved only concepts covered in NAPLAN. They explained that learning and teaching focussed around NAPLAN as not

representative of effective teaching and suggested that it narrowed curriculum which reduced opportunity for pedagogical flexibility and furthermore, stifled student creativity. Teachers reported introducing numerous concepts each day with no opportunity to teach them via appropriate scaffolding or to any real depth and no opportunity for concept consolidation. There was no entry point pertinent to students' cognitive level, 'if students fell behind, so be it'. These concepts were taught in isolation with no links to other concepts or connection to the real world. They reported that the focus on test preparation had been at the expense of other key learning areas. This approach is consistent with New Taylorism where teachers power is usurped and top down power applied.

Consistent with pedagogical practices described by teachers from School A and recognising the limitations of 45 to 60-minute moment-in-time observations as not offering an holistic representation of events, a focus on NAPLAN was observed in all three classrooms visited at School A. This was further confirmed by classroom wall displays, resources and lesson content which revolved around preparation for testing. Lessons involved 'teaching to the test'. Although not observed in Year 7, students provided evidence which confirmed that this had taken place in the two sessions earlier in the day. The Year 7 teacher was observed requesting students call out practice test results in front of their peers; she also displayed records of students' test results on the classroom wall. Pedagogy involving opportunity for higher order thinking, collegiality, student negotiation and movement during learning experiences was only observed during the 45-60-minute observations in Year 7. Focus on practice testing was observed in the form of direct and indirect reference to testing. In one classroom the word 'test' was mentioned 17 times in one lesson. Teacher pedagogy demonstrated at School A was predominantly teacher directed and whole class, involving

lesson instruction by teachers standing alongside whiteboards at the front of rooms. Furthermore, low levels of student engagement, low levels of teacher engagement, and curriculum and pedagogy non-conducive to best teacher practice as described by the Australian Professional Standards for Teachers was also witnessed in classrooms at School A.

In contrast, the implementation of NAPLAN had minimal impact on curriculum, assessment and pedagogical practices at School B. According to School B's teachers, School B focused on student-centered pedagogy and did not introduce any form of alteration to curriculum, pedagogy or assessment practices until two weeks prior to testing. During this time students were introduced to the test format. School B's NAPLAN results were described as above state average and attributed this success to an educational philosophy that involved the whole child and refuted that any further focus on NAPLAN would see an improvement in students' results. They attributed successful student outcomes to formative assessment and reported the use of summative assessment for reporting and planning purposes only.

Teachers at School B reported no changes to their approaches to learning and teaching due to NAPLAN implementation other than that which was unavoidable. Limited test practice which did occur was camouflaged within existing curriculum. Their teaching included a variety of pedagogies such as, individual seat work, whole group instruction and the use of hands-on manipulative materials. They added that basic skills and critical thinking skills were taught and encouraged in their classrooms.

Observations in classrooms at School B supported the account of learning and teaching practices given by the teachers. Wall displays of numeracy and literacy concepts were

shadowed by displays of student achievement. Learning in this environment involved sharing, discovering, problem solving and student negotiation of lesson content and pedagogy. Teachers facilitated while students freely moved from floor to desks when appropriate in rooms where generally there were no whiteboards, blackboards, computers or designated front. There existed an abundance of hands-on manipulative resources and books, both fiction and non-fiction.

NAPLAN practices a positive or negative?

A benefit of NAPLAN preparation identified by School A's teachers was that it reduced student stress as students became more familiar with test structure and process. Furthermore, one teacher claimed that intense test preparation involving constant testing built positive student/teacher relationships based on trust as students were assured that their teacher would not 'let them down'. Observed were disengaged students and teachers 'teaching to the test' in an attempt to improve their results. 'Teaching to the test' and teaching 'test wiseness' made up the lesson content in three of the four lessons observed. This is a further clear example of what Au (2011) warns is a negative impact of high stakes testing regimes, where learning is sacrificed for process.

Those interviewed at School B reported that NAPLAN impacted on learning and teaching only during the two weeks prior to the testing where practice tests were introduced as a way of familiarising students with test structure to ease student anxiety. They did not see the purpose of further time being allocated to test practice and believed that constant practices only improve 'test wiseness' and therefore affected validity of results. Teachers report spending less than two hours a year rehearsing for the tests and believed that constant

practice testing did not improve results. Only in one classroom observed at School B did there appear to be engagement in test preparation. However, this did not include teaching ‘test wiseness’.

National testing equity and validity

School A’s teachers identified the pitfalls of teaching a national testing focused curriculum and identified consequential deficits to students’ education. They were concerned about long-term effects of ‘teaching to the test’ and identified huge gaps in their students’ education. They also raised issues related to Year 3 students’ participation and believed that they were too young to be included.

The length of the testing period was raised by School A’s teachers as a further issue. They reported that expecting students to sit in a test situation for lengthy periods caused them stress and the extensive amount of reading required to complete the tests increased stress for low achieving students when they failed to finish. They added that students still functioning at the concrete operational stage were further disadvantaged due to the absence of time for explicit teaching of concepts using hands-on manipulative materials.

Classroom observations confirmed that the practice of preparation for NAPLAN was inequitable. Witnessed were unattended lower achieving students, and teachers focusing attention on those few students deemed able to succeed. School A’s teachers reported engaging in practices where teaching focused on students below, but extremely close to ‘the line’ indicating the national average, state average or like schools’ average. They reported

disadvantages involved in this practice and related negative consequences to high and low achieving students who also had the right to teacher attention and educational development.

Those interviewed at School B suggested securing equity with regard to NAPLAN testing by ‘levelling the playing field’. They reported that some schools spent an enormous amount of time preparing for NAPLAN and others like School B did not. They questioned who was advantaged or penalised by this practice. Evident during classroom observations was the whole child approach referred to by adults interviewed at School B.

School A’s participants discussed NAPLAN validity and suggested that the opportunity for students to guess successfully or the impact of other influences such as sickness and test anxiety may skew results, rendering them invalid. Observations indicated questionable validity when students engaged in practices involving ‘test wiseness’ and not strategies for problem solving. If the intended purpose of NAPLAN was to gauge students’ ability to pass tests, then this practice was successful. However, if NAPLAN was to gauge students’ academic achievement, results may be invalid as they did not truly reflect data which was directly related to test purpose.

According to School B’s participants NAPLAN did not constitute the best way of collecting valid student data. Regardless of their attempts to ease student stress they reported that students experienced negative effects such as test anxiety. Teachers further suggested that in order to achieve valid data, rigid restrictions related to student preparation, for example teaching to the test and test wiseness, must be implemented. To achieve valid results, they

suggested development of consistency in delivery of curriculum and rigid test guidelines to ensure a level playing field was provided for all students.

Teachers at School A reported using data collected from constant testing to show parents the distance travelled by their children and described this data as more valuable than moment-in-time test data produced by NAPLAN. School B's teachers reported not valuing NAPLAN results and in fact 'gave them very little time' instead placing greater faith in assessment conducted by teachers in classrooms.

Is New Taylorism evident?

This study set out to investigate the experiences of selected teachers at two school sites, exploring the effects on curriculum, assessment and pedagogical practices in their classrooms of differing approaches to NAPLAN. The findings reveal a range of similarities and differences in School A and School B. The school's approach to NAPLAN, and in turn the teachers at these school sites, had a major influence on the curriculum, assessment and pedagogical practices in classrooms. In School A where there was a highly intentional approach to NAPLAN, there was strong harmony with Au's (2011) concept of New Taylorism, where teachers lose pedagogical autonomy as leaders apply top down control to manage the assessment process, with school leaders making decisions and teachers losing autonomy, leading to effects such as narrowing of the curriculum, content drilling and an increase in test focused teaching methods. In School B the effect on the curriculum, assessment and pedagogical practices in classrooms was minimal. However, the sense of political control and an external agenda reinforced the notion of New Taylorism, even in this setting. Au (2011) reminds us through the lens of New Taylorism, high stakes testing regimes

such as NAPLAN may lead to teachers experiencing a sense of increased surveillance, loss of control over curriculum and classroom practice, and a shift of power from teachers to managers. This study confirms that in the two school sites in this study, NAPLAN has had this effect, to varying degrees.

The influence of the school's approach to NAPLAN is clearly evident in these two sites and provides a unique insight into the effects of NAPLAN in schools in Australia where annually more than 1 million students sit the tests. This study adds to the expanding understandings of the effects of NAPLAN in the Australian high-stakes assessment agenda, including the site-specific differences that occur.

Reference List

- Ainsworth, L.B., & Viegut, D. J. (2006). *Common formative assessments: How to connect standards-based instruction and assessment*. Thousand Oaks: Corwin Press. Sage Publications.
- Au, W. (2011). Teaching under the new Taylorism: High-stakes testing and the standardization of the 21st century curriculum. *Journal of Curriculum Studies* 43.1: 25–45.
- Australian Curriculum Assessment and Reporting Authority (ACARA). (2010). National assessment program (NAP): Why NAP? Retrieved from <http://www.nap.edu.au/about/why-nap.html>
- Australian Curriculum Assessment and Reporting Authority (ACARA). (2011). 2011 National Assessment Program Literacy and Numeracy Handbook for principles. Information for parents.

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practise*, 5(1), 7-74.
- Black, P. & Wiliam, D. (2009). To appear *Educational Assessment, Evaluation and Accountability*. *Journal of Personnel Evaluation in Education*. 1(1).
- Broadfoot, P. & Black, P. (2010). *Assessment in Education: Principles, Policy and Practice*. Redefining assessment? The first ten years of assessment in education. *Assessment in Education*, 11(1). DOI: 10.1080/0969594042000208976
- Brown, G. T. L., Irving, S. E., & Keegan, P. J. (2007). *An introduction to educational assessment: Measurement and evaluation*. New Zealand: Pearson Education.
- Caamela, S. (2018). *Management theory of Frederick Taylor*, business.com
Retrieved: 10 November, 2019 <https://www.business.com/articles/management-theory-of-frederick-taylor/>
- Creswell, J. (2007). *Educational research: Planning conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, New Jersey: Pearson Educational Press.
- Cumming, J.J. (2012). *Valuing students with impairment*. Dordrecht: Springer.
- Cumming, J.J, Kimber, K., & Wyatt-Smith, C.M (2011). Historic Australian conceptualisations of English, literacy and multimodality in policy and curriculum and conflicts with educational accountability. *English in Australia*, 46(3), 42-55.
- Donnelly, K. (2010, Aug 02). Flaws test Gillard's proud boast. *The Courier - Mail*. Retrieved from http://edstandards.com.au/index.php?education_standards_institute=103&archive

- Earl, K., & Giles, D. L. (2011). An-other look at assessment. *New Zealand Journal of Teacher's Work*, 8(1), 11-20. Retrieved from www.teacherswork.ac.nz/journal/volume8_issue1/earl.pdf
- Gronlund, N. E., & Waugh, K. W. (2008). *Assessment of student achievement*. Pearson Upper Saddle River, New Jersey: Education, Inc.
- Haladyna, T.M., Downing, S.M., & Rodriguez, M.C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309-334.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning: Tensions and synergies. *The Curriculum Journal*, 16(2), 207-223.
- Harlen, W., & Deakin Crick, R. (2003). Teaching and motivation for learning, *Assessment in Education*, 10(2), 169-208.
- Hardy, I. (2013). Testing that counts: Contesting national literacy assessment policy in complex schooling settings. *Australian Journal of Language and Literacy*, 36(2), 67-78.
- Hohensinn, C., Kubinger, K. D., Reif, M., Schleicher, E., & Khorramdel, L. (2011). Analysing item-position effects due to booklet design within large-scale assessment. *Educational Research and Evaluation*, 17, 497-509.
- Lingard, B., W. Martino, and G. Rezai-Rashti. (2013). Testing regimes, accountabilities and education policy: Commensurate global and national developments. *Journal of Education Policy* 28.5: 539-556.
- Littler, C. (1978). Understanding Taylorism, *The British Journal of Sociology*, Vol. 29, No. 2 pp. 185-202.

- Lobascher, S. (2011). What are the potential impacts of high stakes testing on literacy education in Australia? *Australian Journal of Language and Literacy*, 34(2), 9-19.
- Middle Years of Schooling Association (MYSA) (2008). *MYSA Position Paper: Middle Schooling: People, practices and places*. MYSA: Brisbane.
- Ministerial Council on Education, Employment, Training and Youth Affairs. (MCEETYA). (2008). *Melbourne Declaration on Educational Goals for Young Australians*. Retrieved from www.mceetya.edu.au/mceetya/melbourne_declaration_25979.html
- OECD, (2008). *Assessment for Learning: Formative assessment*, OECD, Paris.
- Plank, S. B. & Condliffe, B. F. (2013). Pressures of the season: An examination of classroom quality and high-stakes accountability. *American Educational Research Journal*, 50(5), 1152-1182. doi: 10.3102/0002831213500691
- Pendergast, D. (2010). Middle schooling. In D. Pendergast, & N. Bahr, (Eds), *Teaching middle years: Rethinking curriculum, pedagogy and assessment (2nded)* (pp. 3-22). Crows Nest, NSW: Allen & Unwin.
- Pendergast, D. & Bahr, N. (2010). *Teaching middle years: Rethinking curriculum, pedagogy and assessment(2nded.)*. Sydney: Allen & Unwin. ISBN 978 1 74237 191 7.
- Polesel, J., Dulfer, N. & Turnbull, M. (2012). *The experience of education: The impacts of high stakes testing on school students and their families – Literature Review*. Sydney, The Whitlam Institute
- Popham, W. (2009). English national curriculum assessment: a commentary from the USA – or exhibiting kindness to the colonies. *Educational Research*, 51, 265-268. doi: 10.1080/00131880902891966

- Rowntree, D. (1987). *Assessing students: How shall we know them?* London, New York: Harper & Row.
- Shepard, L.A. (2003). Reconsidering large-scale assessment to heighten its relevance to learning. In J.M. Atkins, & J.E. Coffey, (Eds.). *Everyday assessment in the science classroom*. Arlington, VA: NSTA Press.
- Stiggins, R. (2005). *Student involved classroom assessment (4th ed.)*. Upper Saddle River, NJ: Prentice Hall.
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. London: Routledge.
- Swain, K (2015). *Middle years students' perceptions and reactions to NAPLAN: The student voice*. (Unpublished doctoral dissertation). Griffith University, Gold Coast.
- Swain, K., & Pendergast, D. (2013). Competing interests: An investigation of National Assessment Program Literacy and Numeracy (NAPLAN) and middle schooling assessment practices. *Australia Journal of Middle Schooling*, 13(1), 4-17
- Swain, K., Pendergast, D. & Cumming, J. (2018). Student experiences of NAPLAN: sharing insights from two school sites. *Australian Educational Researcher*. <https://doi.org/10.1007/s13384-017-0256-5>
- Turan, H. (2015). Taylor's "Scientific Management Principles": Contemporary Issues in Personnel Selection Period, *Journal of Economics, Business and Management*, 3(11), 1102-1105.
- William, D. (2008, September). 'What do you know when you know the test results? The meaning of educational assessments'. Keynote address at the Annual Conference of the International Association for Educational Assessment, Cambridge.

Wyse, D., & W. Torrance. 2009. The development and consequences of national curriculum assessment for primary education in England. *Educational Research*, 51(2), 213-28.

Contact Details:

Doctor Katharine Swain

College of Education, Psychology and Social Work
Flinders University of South Australia
Sturt Road, Bedford Park, SA 5042
Phone: +61 (08) 8201 3421
Mobile: +61(0) 414 453 996
Email: katharine.swain@flinders.edu.au

Professor Donna Pendergast

Head and Dean School of Education and Professional Studies
Griffith University Mt Gravatt Campus
176 Messines Ridge Road
Mt Gravatt Qld 4122
Phone: +61 (07) 3735 1082
Mobile: +61 (0) 466 130 479
Email: d.pendergast@griffith.edu.au