Uses and abuses of standardised testing: Perceptions from high-performing, socially disadvantaged schools

Sarah Penn
Murdoch University, Australia

This study examines the perceptions of teachers and school leaders from high performing, socially disadvantaged government schools about the value of Australia's national standardised test regime for informing learning and teaching. While many studies report negative experiences and perceptions of standardised testing, few studies have explicitly examined the perceptions of high performing schools, especially those that perform higher than expected. Using a qualitative research design, interviews were conducted with teachers (n=12) and principals (n=4) from four schools about their perceptions and experiences of standardised testing. Participants reported that testing was useful for tracking learning trends across years, but was unanimously viewed as an irrelevant tool for measuring the individual achievement of students in their schools. Instead, all schools used regular summative and formative assessments to inform their teaching. The conclusion being that standardised testing is an expensive exercise with little relevance for the stakeholders of this particular educational cohort.

Introduction

Standardised testing is a relatively new practice in Australia, having only been implemented in its current form since 2008. Internationally, particularly in countries such as the United Kingdom and United States, its usage as part of teaching and curriculum practice has been employed for a much longer period of time and "has been implemented with particular intensity" (Au, 2008, p. 639).

Each year, the Australian Government invests millions of dollars on implementing a national standardised testing regime, the National Assessment Program - Literacy and Numeracy or NAPLAN (Carter, Klenowski & Chalmers, 2016). Like other national contexts, general perceptions of standardised testing in Australia are largely negative. Negative perceptions are reported by practitioners and education researchers, as well as parents and the general public (Lee, Lee & Lawton, 2022). NAPLAN has attracted substantial media attention, mostly with negative headlines such as, "How NAPLAN is failing our children" (Bagshaw, 2015), "The NAPLAN curse" (Thomson, 2018) and "NAPLAN test moves exclusively online but critics still want national program scrapped" (Shorthouse, 2022).

NAPLAN's controversy aligns with other arguments towards standardised testing. Throughout the UK, USA and other Western nations, the use of standardised tests has become increasingly widespread. They are used to measure educational outcomes to assist in addressing concerns about the nature and extent of students' learning and accountability behind it (Rizvi & Lingard, 2010). One of the main concerns about standardised testing such as NAPLAN is pressure for schools to perform well on these tests, particularly when test scores are made public. This leads to questions around the

reductive nature of testing when these pressures are in place. Research conducted in Queensland found the following:

> Significant test-centric practices are at play across schools, including, as evidenced in the research presented: A strong focus on teachers meeting, discussing and informing one another about NAPLAN; engaging in curriculum development practices which foreground NAPLAN, and; actively preparing students to sit the test, including, whether intentionally or unintentionally, teaching to the test. (Hardy, 2015, p. 359)

Further concern debated in current literature regarding standardised testing is the impact on the wellbeing of students, teachers and parents alike. According to Dufler, Polesel and Rice (2012), results from their large national survey with over 8,000 responses showed that approximately 90% of teachers surveyed believed that poorer than predicted NAPLAN results would negatively impact a school's reputation, particularly parental perception, the ability to attract and retain students, and staff morale. Klenowski and Wyatt-Smith (2012) noted that "there has been a pervasive silence around the rights of the child/student and the ways in which they have been positioned by testing and accountability priorities" (p. 76).

Regarding NAPLAN's impact and the test-performing abilities between different cohorts of students, Reid (2009) contended that where a student lives is a better indicator of their NAPLAN results than where they go to school. Au (2009) described this phenomenon as the "zipcode effect" where a student's achievement on a test such as NAPLAN correlates with the socio-economic characteristics of where the student lives. Au's research is indicative that this phenomenon is now being observed in Australia.

Other studies have found, however, that NAPLAN is not always associated with negative outcomes and can even have positive impacts. Rogers, Barblett and Robinson (2016) found in their study of 11 independent schools in Western Australia that while emotional distress did rise during and immediately preceding NAPLAN testing periods, it was a very mild rise with no prolonged concerns. Hardy (2014) found that although schools were concerned with improvement of NAPLAN scores for performative purposes, the process provided an opportunity for active engagement on how best to facilitate students' learning. Graham (2016) found NAPLAN and its data were used to drive conversations to promote policy change at the department level in NSW, focusing on other aspects such as attendance, and to shift discourses when considering special needs education.

However, when looking at specific student cohorts such as students situated in schools in low socio-economic status (SES) areas in Australia, research about the capacity of NAPLAN to accurately measure the academic performance of the individual student is limited.

The aim of this paper is to contribute to the literature by examining the perspectives of teachers and principals who teach in low SES schools that have better performance on NAPLAN compared to similar schools. This aim also contributes to the literature about the impact of NAPLAN in low SES communities in particular, as it directly garnered the

opinions of the teachers who use and work with it. The Australian case for standardised testing is interesting from an international perspective because it is firstly, in its infancy compared to standardised testing elsewhere and therefore the teachers and school leaders provide fresh perspectives on its implementation as a relatively new policy initiative. Secondly, within NAPLAN's implementation is the public publishing of school results on an allocated website where stakeholders such as parents can "compare" the results of schools to inform their choice regarding schools for their children. Thirdly, school results are now connected to school funding which operates at a state and federal level with essentially, poorer results equating to more funding (Australian Government Department of Education, Skills & Employment, 2020). This makes NAPLAN very "high-stakes" for schools caught between attracting enrolments with a reputation for sound results published publicly, and results that provide more government funding. These dynamics are quite particular to standardised testing in the Australian context but may hold key lessons for international policymakers and educators observing how these dynamics play out.

Background

Teachers' perceptions of standardised testing are largely negative. A survey of more than 10,000 public school teachers in the US found that only 26% of teachers believed standardised testing was a truthful reflection of what students know (Rebora, 2012). Another survey of US teachers, school leaders and school district administrators found that 73% of participants believed standardised tests are not a good measure of teacher effectiveness, with 72% stating they had access to evidence to support that judgement (Ramirez, Clouse & White Davis, 2014).

Concerns over standardised tests become limiting and the results skewed for students that require variance and choice in the curriculum rather than narrowing (Berliner, 2011) and pedagogies more catered to the cohort (Lewis & Hardy, 2015: Steele, 2014). Essentially, best practice in teaching and pedagogy does not support a "one size fits all" approach and this is especially highlighted within student cohorts that fall outside the median "size" (white, middle class) (Au, 2016; Muller & Boutte, 2019). In Australia, teachers are concerned that standardised testing could have a negative impact on curriculum and pedagogy (Polesel, Dulfer & Turnbull 2012; Ward, 2012), potentially leading to a restricted curriculum that is limited to "teaching to the test" (Cranley, Robinson, Hine & O'Connor, 2022; Hardy, 2015). This latter concern also extends to the "use" of the data after assessment, particularly around teachers' ability to analyse the data effectively to impact practice in the classroom context (Evans, Hatisaru & Williamson, 2021; Raffe & Loughland, 2021).

According to research internationally, standardised testing can have several negative effects on students from low-income and low SES backgrounds. These effects include poor test scores due to inequality in the content and administration of the tests which can highlight systemic racism and the ongoing impact of failure on students who are continually faced with poor results on a test they cannot essentially, score well on (Dianis,

Jackson & Noquera, 2015; Jennings & Sohn, 2014; Tanner, 2015; Trujillo, 2013). Student perceptions of the test can also affect test outcomes, particularly when the explicit purpose of such tests has not been efficiently explained to the student or the teacher does not have a sound understanding of the students' perceptions of the test (Wurf & Povey, 2020).

Standardised testing can also have negative effects on teachers and schools and this concern appears globally (Gonzalez, Peters, Orange & Grigsby, 2017; Smith & Kubacka, 2017; Youn, 2018). Internationally, excessive stress placed on teachers to report good results have caused some teachers to "cheat" by violating testing protocols (Johnson, 2018; Longobardi, Falzetti & Pagliuca, 2018; Richardson, Wheeless & Cunningham, 2008). In Australia, teachers have expressed concern that standardised testing could have a negative impact on staff morale and a school's capacity to attract and retain students (Polesel, Dulfer & Turnbull, 2012).

To counteract the largely negative research on the topic, Buck, Ritter, Jensen and Rose (2010) found while conducting focus groups with US teachers, that the general opinion was positive: "Teachers said the tests provide useful data, the testing regime helps create a road map for instruction, standards and tests don't sap creativity or hinder collaboration, and the accountability imposed by the testing regime is useful" (p. 51). Klein, Zevenbergen and Brown (2006) surveyed 20 elementary, middle and high school teachers and found that some teachers did feel standardised testing provided focus for their teaching.

Many studies have concluded that more research is needed to understand the impacts of standardised testing. In Australia, these calls for more extensive empirical evidence have been especially noted, given the high-stakes nature of NAPLAN (Rose, Low-Choy & Singh, 2020). Polesel et al (2012, p. 9) noted that "further research is required to examine carefully the uses, effects and impacts of NAPLAN, as reported by a range of users, including systems, the teaching profession, parents and students." Rogers, Barblett and Robinson (2016, p. 340) concluded that "a lack of research studies investigating the effect of NAPLAN on stakeholders currently limits arguments about NAPLAN's impact."

Australian context

Since 2008, all students attending school in Australia in Years 3 and 5 (primary school) and 7 and 9 (secondary school) are assessed in the areas of reading, writing, language conventions (spelling, grammar and punctuation) and numeracy by NAPLAN. The tests are undertaken in May of every year by students in the aforementioned year levels and are administered by the Australian Curriculum, Assessment and Reporting Authority (ACARA), an independent statutory authority.

NAPLAN results are disseminated in several ways. Students and their caregivers receive a report of the student's results. Individual student results are also provided to teachers and school leaders. Departments of Education in each jurisdiction (state/territory) also receive

results for all students (and schools) in their jurisdiction. Finally, ACARA publishes average results, by test domain and year level, for every school on the publicly available *MySchool* website. Each school's NAPLAN performance is compared with the national average as well as with "like" schools, which are schools that have similar levels of socio-educational (dis)advantage based on student SES status, Indigenous background, and school location.

The official purpose of NAPLAN is quoted similarly on the NAPLAN, ACARA and *MySchool* websites as the following: "It is a nationwide measure to see whether or not young Australians are developing the literacy and numeracy skills that provide the critical foundation for other learning and for their productive and rewarding participation in the community." (ACARA, 2023a). According to the NAPLAN parent information (ACARA, 2023b), found on the aforementioned websites, the NAPLAN results are used for the following:

- Students and parents may use individual results to discuss progress with teachers.
- Teachers use results to help identify students who need greater challenges or extra support.
- Schools use results to identify strengths and areas of need to improve teaching programs, and to set goals in literacy and numeracy.
- School systems use results to review the effectiveness of programs and support offered to schools.
- The community can see information about the performance of schools over time at myschool.edu.au (ACARA, 2023b)

However, there is little detail regarding policy initiatives to improve or increase these skills for young Australians in order to ensure test results continue on an upward trajectory, which currently, they are not, with National preliminary figures for 2019 showing a majority of stagnant scores with a decline in literacy and numeracy levels over the past few years. (Turner & Pale, 2019).

Approach

An exploratory qualitative research design within an interpretive paradigm was used to examine the understandings and experiences of teachers' pedagogical practice and the feelings and beliefs associated with them. Random purposive sampling and data from the Australian Government's *MySchool* website was used to identify socially disadvantaged government primary schools whose performance on NAPLAN was substantially greater than schools with similar socioeconomic compositions. As in-depth interviews were being conducted, the study was limited to the state of Victoria, where the author was living at the time of data collection. Victoria is Australia's second most populous state, and its capital Melbourne is Australia's second largest city. Seven schools met the selection criteria and four schools agreed to participate in the study. Table 1 presents school profiles at the time of data collection.

Table 1: Participant schools' profiles from *MySchool* (website https://www.myschool.edu.au/)

	School A	School B	School C	School D
School sector	Government	Government	Government	Government
Year range	Prep-6	Prep-6	Prep-6	Prep-6
Location	Inner regional	Outer regional	Inner regional	Major cities
Enrolments	99	203	108	454
	Boys 52	Boys 103	Boys 51	Boys 221
	Girls 47	Girls 100	Girls 57	Girls 233
% Indigenous	9%	37%	13%	3%
% Language background other than English	4%	2%	12%	3%
School ICSEA value (average Australian value – 1000)	915	853	823	992
Distribution of students according to socio educational advantage quarters				
Top quarter	2%	3%	1%	12%
Middle quarters	40%	31%	10%	54%
Bottom quarter	58%	66%	89%	34%

Participants included the principal and three teachers from each school. Principals nominated teachers who they believed were exemplary in their practice. The principals of each school had several attributes in common. All four principals had over 25 years' experience working in education, with two principals having 36 years' experience. All four principals had been teachers previously and all four had taught in small, regional schools at some point during their career. All four principals had a range of teaching experiences. Their ages ranged from 48 to 62 years, with two principals being male and two being female. With respect to their personal backgrounds, two principals identified as middle-class and two as working class. The two principals that identified as working-class were still living in the same areas they were raised in.

The 12 teacher-participants (10 female and two male) ranged in teaching experience. The least experienced teacher had been teaching for four years, all at their current school. Four teachers had between 10 and 20 years' experience, two teacher-participants with between 20 and 30 years' experience, four teachers with between 30 and 40 years' experience and one teacher with just over 40 years' experience in teaching. Similarly, all but one teacher-participant had spent most of their teaching career at their present school, with most having spent between 16 and 30 years teaching at their current school. As a result, every teacher stated that the bulk of their experience, if not all, had been working with students from low SES backgrounds and poverty. The age range of the teacher-participants was between 27 and 62 years.

Data collection comprised in-depth individual interviews with preparation, permissions, interviews and collation spanning two years ending with analysis completed mid-2018. As the study focused on participants' experiences and perceptions of NAPLAN, interviews were conducted to "capture the deep meaning of experience in the participants' own words." (Marshall & Rossman, 2011, p. 93). Each interview was allocated one hour,

though most went considerably over this time allocation. The interviews were recorded. The interview schedule regarding NAPLAN centred on the following questions:

- As a teacher, what are your thoughts and views on NAPLAN and particularly on its use as an assessment tool for grading students and schools?
  - What are the advantages?
  - What are the disadvantages?
- In what ways has NAPLAN impacted your students' learning?
  - What are the advantages and disadvantages?

The recorded interviews were transcribed and then analysed using interpretive and thematic analysis. Following the guidelines of Saldana (2016), transcripts were coded for key words and insights, then the codes were grouped into larger categories and sub-categories. Overarching themes were then identified.

Findings

Based on the previous review of the literature, largely negative perceptions of NAPLAN were expected and standardised testing more generally. However, the participants were circumspect and pragmatic in their opinions. All the participants explained the individualised testing for each student at regular intervals throughout the year and in the case of two schools, time and funding was devoted to this process. For example, one school in this dataset funds relief teachers to cover classroom teachers at a cost of approximately $10,000 annually to allow classroom teachers to assess and review students individually every five to six weeks. The justification for this was that it allows for a more targeted approach for each individual student, ensuring they continue on an "upwards trajectory" with their learning. The three major themes that emerged from the analysis were *Frustration and fear*, *Masking* and *A relevant use*.

Theme one: Frustration and fear

This theme encompassed nearly all of the negative attributes associated with NAPLAN as a form of standardised testing. In the first instance, the participants expressed frustration, not for the test itself but how it is viewed by the stakeholders, namely the students, parents and society at large and how it is inefficiently used. One principal from School D illustrated his point with an example from his own life:

> I'll use my daughter as an example, she was dux at school, very, very clever girl but in her grade 3 NAPLAN she got her results and she opened her envelope because she's a smart little cookie and pulled it out and stuffed it down in the bottom of the bag because with a quick look at the graph she thought she was underachieving and at the bottom of the pile instead of being in the top 5% of Australia so she hid it underneath her jumper and everything. She was devastated. Can you imagine that little kid in grade 3 who has never achieved anything takes that NAPLAN test home to Mum and there you are in the bottom 10% of Australia? So it's not the top 10% that we're worried about, it's the effect that has on a little grade 3 kid. That dent in their self-esteem is very hard to recover

> from. They know they're a dummy right from that NAPLAN result coming in. They stereotype themselves and that's what we're fighting against now, trying to break the kids through that barrier and it's really bloody difficult once that first NAPLAN test is done and it goes home to parents. So that is something that I think is really, really detrimental to kids' future achievements. (Principal, Participant 13)

There was a consensus amongst all the participants bar two, that this frustration of how NAPLAN is viewed by the wider community can make it difficult to relay to their students, who may already have resilience and self-esteem issues that it is not reflective of everything they can do. They were concerned about the negative impact this can have on a student's personal belief in themselves as a learner and the ongoing impact that may have on their school life.

> I think too much stock is put into NAPLAN. It's a group of tests in one week way too early in the year, which can set the kids up for failure and I don't believe it is an accurate portrayal of what those students can do and how far those students have come. It is purely for the governments and stakeholders so I put very little bearing on it. For example, my kids are now at a six, which is at the end of prep level, so as a teacher I do not look like I've done my job, yet I know those kids have come a long way from when they arrived in my classroom or indeed when they arrived at the school. (Teacher, Participant 10)

Several participants expressed frustration at how NAPLAN was unable to show children's improvement and did not take into account issues that students may experience academically or out of school.

> Disadvantages [of NAPLAN] can be it's a bit frustrating for the staff because some of the students don't test well and then it's harder with that resilience and it could be down to they can't sit there for that long or they can't read the instructions because their reading abilities don't support them like what we do to support them. It's not normally how they work. So, I think that frustration for teachers is it doesn't capture all the other improvements the kids have made so that's probably the biggest disadvantage that it doesn't necessarily reflect what the teachers are working on. (Principal, Participant 9)

> Often our kids have issues at home so NAPLAN day they may have been up till three in the morning, they might have had a huge fight in the family before coming to school. Our kids tend to have more bad mornings that we have to salvage a day from than good mornings and great days but NAPLAN can't take that into account. (Teacher, Participant 11)

Here lies the frustration that the participants expressed for their students. NAPLAN can create issues of inadequacy or failure for students who may not understand that the assessment is a test that does not coherently reflect what they may be achieving in the classroom.

Theme two: Masking

Several participants discussed their perceptions that some other (i.e., not their own) schools and teachers are "masking" their results by employing a range of strategies to make their NAPLAN results appear different to what they are. Teaching to the test, barring students who are deemed to lower the result average for the school in completing NAPLAN and the merging of school results in regional areas are the main concerns identified by the participants. The participants who discussed this issue, indicated that this masking behaviour was widespread and were concerned that NAPLAN results for schools that participated in this practice were inaccurate. The following was discussed in regard to "Teaching to the test":

> At one stage, the grade 2's were [previously at participant's own school] being taught results from previous NAPLAN tests to prepare them for the questions. That so annoyed me as that was valuable teaching time being used. Now all these years later all pre-service teachers know is how to administer NAPLAN, they should be being taught how to test and monitor the children themselves rather than relying on results from NAPLAN. Everyone knows teaching to the test is a big no. (Teacher, Participant 16)

Four participants described a second aspect of masking, which was barring or withdrawing students that may lower a school's NAPLAN average.

> We know that kids at other schools are told they don't need to come that day or are gathered up and taken to the library to do an activity while NAPLAN is administered, not good. (Teacher, Participant 8)

The third aspect of masking related to regional schools in the same area having their NAPLAN results merged and combined and as illustrated in an example below:

> The schools who are worrying about that (what is in the test) are worrying because they can't teach and they are worried about what it might unmask. For example, ********* **** is a terrible school, the teachers are all dropouts and the students suffer but you can't get their NAPLAN results because it merges with *******'s so there is a lot of masking going on. (Principal, Participant 5)

In essence, this theme of masking involved concerns surrounding schools attempting to doctor NAPLAN results by using strategies such as specifically "training" students to sit NAPLAN tests, omitting students who may lower the school average result, and lower performing schools being merged with other schools in their locale. These strategies were deemed as "cheating" by the participants who discussed this theme.

Theme three: A relevant use

While critical of NAPLAN, participants did articulate relevant uses of the test in their own practice and for their school. This ranged from comparing growth between year levels (Year 3 and Year 5) and highlighting any areas that require further instruction.

> We use NAPLAN to look at the movement between Years 3 and 5. We also look at growth to have happened right across the board not just at the tail end so that all children are still being challenged and are moving on and it is great indicator for that purpose. You can also check on anomalies that appear in the data and use it for reflection. There is no big build up with the test. So it is just a reflective tool for us and it works well in that regard. (Teacher, Participant 6)

> I always print off and match the NAPLAN cohort data and compare it to state as well. We do really well and it's an opportunity to go back to staff and celebrate. There's years where we don't do as well as others but generally it shows what we are doing is working and again, if it is not as positive as we like that is why we have the school data and I have that for my region as well. (Principal, Participant 9)

> We will send out when it's coming up to NAPLAN time and we explain that it is a one-off test and that a student might perform better or worse on the day. So individually I think it can be very wrong but collectively I don't think it is. So Sally* might do worse on the day and Dave* might do better but overall the kids will collectively be on the mark. We have a rule here that if a kid performed better than what was expected or performed worse than what was expected we get the parents in and talk to them. We have an assessment schedule so the teachers know where each child is at so we talk about NAPLAN in regard to the assessment schedule. (Principal, Participant 5) *Names changed.

> So the great thing about NAPLAN is that we've been able to use it to analyse what we're teaching here against similar cohorts of schools and children and we can see where we're failing. Not so much failing but not actually achieving the results we should be. So we are comparing similar cohorts and it's great for mapping continual growth between grade 3 and grade 5 so we can evaluate our programs from grade 3, 4 and 5 and prior to that we can evaluate our programs in the infant areas to see whether or not we're actually hitting our targets. (Principal, Participant 13)

Therefore, all four schools did use NAPLAN data to inform their direction in learning. However, all stated that they saw it as a secondary or supplementary tool to their own assessment schedules for the students. Nearly all participants discussed that due to the large variation in academic ability within their classrooms and schools, assessment schedules were individualised for every student so achievement could be tracked on a more individualised and therefore, more truthful level targeted to each student. One Principal participant stated that

> … we should be able to have access to those results (NAPLAN) within a couple of minutes, not 3 or 4 months later … providing you know how to use that information and how to access that information is fantastic but 3 months down the track – what's the point?

Summary of findings

Participants expressed their concerns with NAPLAN and the considerations when administering a standardised test to this particular cohort of students. However, benefits from the data produced by the test were discussed but very much on a school more than

an individual student level and as a secondary assessment tool to each school's own assessment schedules and processes.

Discussion

The participants' assessments of NAPLAN were very pragmatic and more positive than expected based on the findings of other studies. While participants did discuss concerns that were directed at the behaviours of other schools when administering and reporting on NAPLAN, when it came to their own students, the discussion was very matter of fact. As all of the schools interviewed completed rigorous individual testing of their students and used this to drive their programming for the classroom, the general view of NAPLAN is that it is one test on one day and as one principal stated, "doesn't get a lot of air play." Some participants mentioned the merit of using it as trend data between Years 3 and 5. All the schools discussed it as a general overview of where their students are at academically and for this purpose, it held merit for reflection of what their own data and assessment was showing. Then if the NAPLAN data was very different from their own, there was opportunity to take a closer look at what NAPLAN was showing. However, all the principal participants said this rarely happened and generally it was on par with their own, so NAPLAN data was never viewed any deeper than a reflection on what their own data was showing.

Pragmatism regarding NAPLAN was prevalent throughout the data for this research study. The teachers and principals downplay NAPLAN in their schools as the students that they are working with are unlikely to have strong skills in "test sitting" anyway, which is a big requirement of being successful on a NAPLAN test. As one participant stated,

> I know personally I never performed well in tests as I have test anxiety and with these children they can feel that pressure quite strongly so even if they are quite good in a certain subject area, this may not be relayed in the test and that has to be remembered. (Teacher, Participant 10)

There tended to be a general consensus from the participants that NAPLAN is not designed with their particular cohort of students in mind and therefore the data that it conveys is to be taken "with a pinch of salt." Therefore, while the participants were pragmatic, there certainly wasn't the overwhelming negative feeling that was hypothesised for this study. However, there was a sense of containment or downplay so that the students would not be too affected either in their school routine or in their final results. The considerations for this will be discussed in the following section.

Implications

According to ACARA (quoted in Carter, Klenowski & Chalmers, 2016) the annual costs in relation to NAPLAN testing are approximately between 7 and 7.5 million Australian dollars per year with the government costs of administering the tests between $19 and $21 per student. Jen Buchanan of Future Schools puts the figure at $45 per student (Buchanan, 2020). When funding is so tight for schools working with students from low

SES backgrounds, the hope would be that NAPLAN is a powerful "measure through which governments, education authorities, schools and the community can determine whether or not young Australians are meeting important educational outcomes" (ACARA, 2022). As mentioned previously, there is no information around policy change if the above is not occurring. The data from this research study shows that NAPLAN is not viewed as a particularly relevant measure for students from low SES backgrounds.

The inference from this is that stakeholders working in low SES areas with a particular cohort of students are being measured against criteria/curriculum outcomes that students in these schools may not have reached yet or contain content that students may not have learnt yet. All the participants stated that there is a much wider spectrum of ability when their students enter school and indeed, at each year level and this can have a strong effect on a school's results in a standardised test such as NAPLAN. With the move to NAPLAN online, the tailored testing model that comes with this change will allow the test to adapt in real time to the student's perceived ability level, which will assist with student engagement and individualised test scoring but is unlikely to positively affect the school band placement for results. The schools in this study already participate in rigorous assessment of students on a very individualised basis and this allows them to monitor and cater for each individual student across an often vast range of ability levels.

While the participants in the study were not dramatically against NAPLAN, their opinions ranged from not relevant to somewhat interesting for looking at trend data between Years 3 and 5. All four schools showed very little concern about how their results are displayed on the *MySchool* website and showed no interest in comparing themselves to "like" schools. Participants from all four schools stated they couldn't waste the time in class as the teaching time they have with their students is too valuable to lose on teaching "test-sitting." The schools all showed a lack of interest in the competitive aspect of NAPLAN which has been proven in the literature to cause varying levels of stress for other schools, teachers and students. According to the *MySchool* website, the schools in this study had very few "competitor" schools in their own state (Victoria). This would suggest that NAPLAN has a less negative impact on schools and teachers in non-competitive education marketplaces.

The question to be asked then, is NAPLAN worth it? For the costs and the time, it incurs, the data in this research study suggests that it isn't – at least not for the schools themselves. When these schools need funding in so many other areas that could contribute to student academic success, a test that provides some minor trend data and a small opportunity to give themselves a pat on the back does not seem a valid and sufficient use of funds and time. Indeed, one could deduce from the participants' responses that they are successful in their NAPLAN scores because ironically, they pay very little attention to it. However, three of the principal participants did discuss that they were aware of schools who were very concerned about NAPLAN results and in their attempts to implement strategies of concern mentioned previously, either performed poorly or masked the fact that the students of that school were not academically succeeding. This has very strong implications for the purpose of NAPLAN moving into the future.

Two areas of concern were identified when it came to NAPLAN and other forms of standardised testing from the literature. Firstly, there was a question around the relevance of NAPLAN and the data it produces for academic assessment. This is particularly poignant for students from low SES backgrounds where research has previously indicated that a student's results on a test like NAPLAN correlates with the socio-economic area where that student lives. (Au, 2009; Lange & Meaney, 2011; Thompson & Harbaugh, 2013) The second concern was centred on the impact on the wellbeing of students, teachers and parents alike. The hypothesis for this aim was for a general consensus that the participants of this research study would see NAPLAN as highly problematic and an unreliable indicator of a school's academic success.

The responses from the participants supported the hypothesis…to an extent. The underlying assumption was that the hypotheses would stem from a general view that NAPLAN is a cause of stress and pressure for schools, teachers and students alike. This was not the case for the schools in this research study. These schools all had very structured data collection and individualised testing already in place, so NAPLAN was something they do as a mandatory requirement. The results of NAPLAN were only relevant as secondary data to confirm the data they were already collecting within their schools. There was also a general acceptance that NAPLAN was somewhat irrelevant as it did not capture what their particular cohort of students could really do. What was a fascinating conclusion from this section of the research study was that ironically, the nonchalant view of NAPLAN and the very little emphasis that was placed on it within the schools may have actually contributed to these schools scoring well on the test.

This research study, whilst in depth, used a very small sample of schools and education professionals to capture opinions on the topic of NAPLAN. Further research would clarify NAPLAN's worth as an accurate measurement for schools with cohorts of students from low SES backgrounds. Comparatively, future research on similar schools that are "low-performing" on NAPLAN would provide important clarification into the extent of the efficacy of the strategies discussed here and the extent of the effect of the themes and issues raised in this research. The results of such research could lead to the identification of systemic issues and the collaboration of "higher-performing" schools sharing their philosophies and strategies to assist other schools. (Buchanan & Hellstén, 2013)

Conclusion

Whether positive or negative, enlightening or subversive, NAPLAN remains the only major measurement tool available in Australia for stakeholders such as parents, teachers, policymakers and researchers to use for the purposes of data review, association, or collation of educational achievement. Therefore, having the opportunity to use NAPLAN to select study participants and then being able to discuss the validity of this process with one of the major stakeholders of this data, teachers, is an opportunity to contribute to the literature using the voices of the major players. Teachers are the players in this game who deliver, administer and observe the outcomes of NAPLAN (and not just the educational

outcomes) and are often under the most scrutiny when results are made public (Fogelgarn & Burns, 2020; Shine & Rogers, 2021; Thompson, 2013).

The teachers interviewed for this study work with students specifically from low socio-economic backgrounds. This allows provision of further insight into how this even more focussed group of stakeholders are coping with the integration of NAPLAN into their teaching environment. This point becomes particularly poignant when current research indicates that the lowest achieving students on the NAPLAN scales are from disadvantaged backgrounds and include students who identify as having disabilities, are Indigenous and/or come from low socio-economic backgrounds (Lange & Meaney, 2011; Thompson & Harbaugh, 2013). To abridge the findings of this study, it concludes by quoting Spolsky (quoted in Mac Ruairc, 2009, p. 47) who said:

> Tests should be labelled just like dangerous drugs: 'Use with care!'

Acknowledgements

The findings of this study formed part of a larger study taken from the thesis of the author: Penn, S. A. (2019). *"How do they do it?" An investigation into the practices of Australian teachers and principals working in low SES schools*. Professional doctorate thesis, Murdoch University. https://researchrepository.murdoch.edu.au/id/eprint/53648/

The author would like to acknowledge Associate Professor Laura Perry and Mrs Lorraine Jacob (Murdoch University) for their assistance in the first drafts of this article.

References

ACARA (Australian Curriculum, Assessment and Reporting Authority) (2022). *Why NAP?* http://www.nap.edu.au/about/why-nap

ACARA (2023a). *NAPLAN explained*. https://www.myschool.edu.au/naplan-explained

ACARA (2023b). *Information for NAPLAN parents and carers 2023*. https://nap.edu.au/docs/default-source/default-document-library/naplan-information-brochure-for-parents-and-carers.pdf

Au, W.W. (2008). Devising inequality: A Bernsteinian analysis of high-stakes testing and social reproduction in education. *British Journal of Sociology in Education,* 29(6), 639-651. https://www.jstor.org/stable/40375388

Au, W. (2009). *Unequal by design: High-stakes testing and the standardization of inequality*. New York: Routledge. [2nd ed., 2022] https://www.routledge.com/Unequal-By-Design-High-Stakes-Testing-and-the-Standardization-of-Inequality/Au/p/book/9780367437039

Au, W. (2016). Meritocracy 2.0: High-stakes, standardized testing as a racial project of neoliberal multiculturalism. *Educational Policy,* 30(1), 39-62. https://doi.org/10.1177/0895904815614916

Australian Government Department of Education, Skills & Employment. (2020). What is the Schooling Resource Standard and how does it work? [see 23 February version at https://www.education.gov.au/recurrent-funding-schools/schooling-resource-standard]

Bagshaw, E. (2015). How NAPLAN is failing children. *The Sydney Morning Herald*, 25 September, p.11.

Berliner, B. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education,* 41(3), 287-302. https://doi.org/10.1080/0305764X.2011.607151

Buchanan, Jen (2020). Standardised testing… is it worth it? *Future Schools.* https://futureschools.education/

Buchanan, John & Hellstén, M. (2013). Editorial: "I love the smell of NAPLAN in the morning". *Issues in Educational Research,* 23(3), ii-iii. http://www.iier.org.au/iier23/editorial23-3.pdf

Buck, S., Ritter, G. W., Jensen, N. C., & Rose, C. P. (2010). Teachers say the most interesting things - an alternative view of testing. *Phi Delta Kappan,* 91(6), 50-54. https://doi.org/10.1177/003172171009100613

Carter, M. G., Klenowski, V. & Chalmers, C. (2016). Who pays for standardised testing? A cost-benefit study of mandated testing in three Queensland secondary schools. *Journal of Education Policy,* 31(3), 330-342. https://doi.org/10.1080/02680939.2015.1070206

Cranley, L., Robinson, C., Hine, G. & O'Connor, D. (2022). The desks have changed; it must be NAPLAN time: How NAPLAN affects teaching and learning of mathematics. *Issues in Educational Research,* 32(4), 1306-1320. http://www.iier.org.au/iier32/cranley.pdf

Dianis, J. B., Jackson, J. H. & Noguera, P. (2015). High-stakes testing hasn't brought education gains. *Phi Delta Kappan,* 97(1), 35-35. https://doi.org/10.1177/0031721715602235

Dufler, N., Polesel, J. & Rice, S. (2012). *The experience of education: The impacts of high stakes testing on school students and their families.* The Whitlam Institute, University of Western Sydney, Australia. https://core.ac.uk/download/pdf/30677832.pdf

Evans, D., Hatisaru, V. & Williamson, J. (2021). The use of NAPLAN data and support for it: Perceptions of practising teachers. *Australian Educational Leader,* 43(3), 69-74. https://search.informit.org/doi/abs/10.3316/INFORMIT.115090195898219

Fogelgarn, R. & Burns, E. A. (2020). What constrains passionate teaching? A heuristic exploration. *Issues in Educational Research,* 30(2), 493-511. https://www.iier.org.au/iier30/fogelgarn.pdf

Gonzalez, A., Peters, M. L., Orange, A. & Grigsby, B. (2017). The influence of high-stakes testing on teacher self-efficacy and job-related stress. *Cambridge Journal of Education,* 47(4), 513-531. https://doi.org/10.1080/0305764X.2016.1214237

Graham, L. J. (2016). Reconceptualising inclusion as participation: Neoliberal buck-passing or strategic by-passing? *Discourse: Studies in the Cultural Politics of Education,* 37(4), 563-581. https://doi.org/10.1080/01596306.2015.1073021

Hardy, I. (2014). A logic of appropriation: Enacting national testing (NAPLAN) in Australia. *Journal of Education Policy,* 29(1), 1-18. https://doi.org/10.1080/02680939.2013.782425

Hardy, I. (2015). A logic of enumeration: The nature and effects of national literacy and numeracy testing in Australia. *Journal of Education Policy,* 30(3), 335-362. https://www.tandfonline.com/doi/abs/10.1080/02680939.2014.945964

Jennings, J. & Sohn, H. (2014). Measure for measure: How proficiency-based accountability systems affect inequality in academic achievement. *Sociology of Education,* 87(2), 125-141. https://doi.org/10.1177/0038040714525787

Johnson, L. (2018). "Advocacy cheating": Teacher noncompliance in standardised literacy testing. *English Teaching: Practice & Critique,* 17(3), 168-181. https://doi.org/10.1108/ETPC-11-2017-0160

Klein, A. M., Zevenbergen, A. A. & Brown, N. (2006). Managing standardized testing in today's schools. *The Journal of Educational Thought / Revue de la Pensée Éducative,* 40(2), 145-157.

Klenowski, V. & Wyatt-Smith, C. (2012). The impact of high stakes testing: The Australian story. *Assessment in Education: Principles, Policy & Practice,* 19(1), 65-79. https://doi.org/10.1080/0969594X.2011.592972

Lange, T. & Meaney, T. (2011). I actually started to scream: Emotional and mathematical trauma from doing school mathematics homework. *Educational Studies in Mathematics,* 77(1), 35-51. https://doi.org/10.1007/s10649-011-9298-1

Lee, J., Lee, J.-S. & Lawton, J. (2022). Cognitive mechanisms for the formation of public perception about national testing: A case of NAPLAN in Australia. *Educational Assessment Evaluation and Accountability,* 34, 427-457. https://doi.org/10.1007/s11092-022-09398-8

Lewis, S. & Hardy, I. (2015). Funding, reputation and targets: The discursive logics of high-stakes testing. *Cambridge Journal of Education,* 45(2), 245-264. https://doi.org/10.1080/0305764X.2014.936826

Longobardi, S., Falzetti, P. & Pagliuca, M. M. (2018). Quis custiodet ipsos custodes? How to detect and correct teacher cheating in Italian student data. *Statistical Methods & Applications,* 27(3), 515-543. https://doi.org/10.1007/s10260-018-0426-2

Mac Ruairc, G. (2009). 'Dip, dip, sky blue, who's it? NOT YOU': Children's experiences of standardised testing: A socio-cultural analysis. *Irish Educational Studies,* 28(1), 47-66. https://doi.org/10.1080/03323310802597325

Marshall, C. & Rossman, G. B. (2011). *Designing qualitative research*. Thousand Oaks, CA: SAGE. [7th ed.] https://us.sagepub.com/en-us/nam/designing-qualitative-research/book274291

Muller, M. & Boutte, G. S. (2019). A framework for helping teachers interrupt oppression in their classrooms. *Journal for Multicultural Education,* 13(1), 94-105. https://doi.org/10.1108/JME-09-2017-0052

Polesel, J., Dulfer, N. & Rice, S. (2012). *The experience of education: The impacts of high stakes testing on school students and their families: An educator's perspective.* Sydney: Whitlam Institute within the University of Western Sydney. https://www.whitlam.org/publications/2017/10/17/the-experience-of-education-the-impacts-of-high-stakes-testing-on-school-students-and-their-families

Raffe, C. P. & Loughland, T. (2021). "We're not data analysts": Teachers' perspectives on factors impacting their use of student assessment data. *Issues in Educational Research,* 31(1), 224-240. http://www.iier.org.au/iier31/raffe.pdf

Ramirez, A., Clouse, W. & White Davis, K. (2014). Teacher evaluation in Colorado: How policy frustrates practice. *Management in Education,* 28(2), 44-51. https://doi.org/10.1177/0892020613511264

Rebora, A. (2012). Teachers place little value on standardized testing. *Education Week,* 27 March. https://www.edweek.org/teaching-learning/teachers-place-little-value-on-standardized-testing/2012/03

Reid, A. (2009). Is this a revolution? A critical analysis of the Rudd government's national education agenda. *Curriculum Perspectives*, 29(3), 1-13.

Richardson, B. K., Wheeless, L. R. & Cunningham, C. (2008). Tattling on the teacher: A study of factors influencing peer reporting of teachers who violate standardized testing protocol. *Communication Studies,* 59(3), 202-219. https://doi.org/10.1080/10510970802257531

Rizvi, F. & Lingard, B. (2009). *Globalizing education policy*. London: Routledge. https://www.routledge.com/Globalizing-Education-Policy/Rizvi-Lingard/p/book/9780415416276

Rogers, S. L., Barblett, L. & Robinson, K. (2016). Investigating the impact of NAPLAN on student, parent and teacher emotional distress in independent schools. *The Australian Educational Researcher,* 43(3), 327-343. https://doi.org/10.1007/s13384-016-0203-x

Rose, J., Low-Choy, S., Singh, P. & Vasco, D. (2020). NAPLAN discourses: A systematic review after the first decade. *Discourse: Studies in the Cultural Politics of Education,* 41(6), 871-886.

Saldana, J. (2016). *The coding manual for qualitative researchers*. Thousand Oaks, CA: SAGE. https://us.sagepub.com/en-us/nam/book/coding-manual-qualitative-researchers-1

Scot, T. P., Callahan, C. M. & Urquhart, J. (2009). Paint-by-number teachers and cookie-cutter students: The unintended effects of high-stakes testing on the education of gifted students. *Roeper Review,* 31(1), 40-52. https://doi.org/10.1080/02783190802527364

Shine, K. & Rogers, S. L. (2021). Parents and teachers' perceptions of the tone and emotional impact of education news coverage. *Journalism and Media,* 2(2), 193-207. https://doi.org/10.3390/journalmedia2020011

Shorthouse, J. (2022). NAPLAN test moves exclusively online but critics still want national program scrapped. *ABC News,* 3 May. https://www.abc.net.au/news/2022-05-03/naplan-online-adaptive-testing-critics-call-to-scrap-program/101032924

Smith, W. C. & Kubacka, K. (2017). The emphasis of student test scores in teacher appraisal systems. *Education Policy Analysis Archives,* 25(86), 1-29. https://doi.org/10.14507/epaa.25.2889

Steele, L. A. (2014). Peddling pedagogies: The winners and losers of a standardized testing economy. *Radical Teacher,* 100, 153-155. https://doi.org/10.5195/rt.2014.123

Tanner, D. (2015). Sorry, President Obama, but high-stakes tests can't cure cancer. *Kappa Delta Pi Record,* 51(1), 4-6. https://doi.org/10.1080/00228958.2015.988563

Thompson, G. (2013). NAPLAN, MySchool and accountability: Teacher perceptions of the effects of testing. *International Education Journal: Comparative Perspectives*, 12(2), 62-84. https://files.eric.ed.gov/fulltext/EJ1017709.pdf

Thompson, G. & Harbaugh, A. G. (2013). A preliminary analysis of teacher perceptions of the effects of NAPLAN on pedagogy and curriculum. *The Australian Educational Researcher,* 40(3), 299-314. https://doi.org/10.1007/s13384-013-0093-0

Thomson, K. (2018). The NAPLAN curse - Australian schools in crisis, anxious students and stressed-out teachers. *60 Minutes,* 5 August. https://www.9now.com.au/60-minutes/2018/episode-23

Trujillo, T. M. (2013). The disproportionate erosion of local control: Urban school boards, high-stakes accountability, and democracy. *Educational Policy,* 27(2), 334-359. https://doi.org/10.1177/0895904812465118

Turner, K. & Pale, M. (2019). The relationship between student attitudes to school survey results and NAPLAN results. *Issues in Educational Research,* 29(1), 282-300. http://www.iier.org.au/iier29/turner.pdf

Ward, D. M. (2012). *The effects of standardised assessment (NAPLAN) on teacher pedagogy at two Queensland schools.* PhD thesis, Queensland University of Technology. http://eprints.qut.edu.au/63662/

Wurf, G. & Povey. R. (2020). "They show how smart you are": A mixed methods study of primary students' perceptions of assessment tasks. *Issues in Educational Research,* 30(3), 1162-1182. http://www.iier.org.au/iier30/wurf.pdf

Youn, M. (2018). The influence of standardized testing pressure on teachers' working environment. *KEDI Journal of Educational Policy,* 15(2), 3-22. https://www.kedi.re.kr/eng/kedi/bbs/B0000005/view.do?nttId=719&menuNo=200067&divisionCd=35&pageIndex=1

Dr Sarah Penn *BA/BT MA MEd EdD*, School of Education, Murdoch University, Australia, is an award-winning educator who has been working in early childhood, primary and vocational education for 18 years, both nationally and internationally. Her research interests include early childhood education and equity in education.
Email: sarahapenn@hotmail.com

Please cite as: Penn, S. (2023). Uses and abuses of standardised testing: Perceptions from high-performing, socially disadvantaged schools. *Issues in Educational Research*, 33(1), 266-283. http://www.iier.org.au/iier33/penn.pdf