



Queensland University of Technology
Brisbane Australia

This is the author's version of a work that was submitted/accepted for publication in the following source:

[Carter, Marilyn Gladys, Klenowski, Valentina, & Chalmers, Christina](#)
(2016)

Who pays for standardised testing? A cost-benefit study of mandated testing in three Queensland secondary schools.

Journal of Education Policy, 31(3), pp. 330-342.

This file was downloaded from: <https://eprints.qut.edu.au/86605/>

© Copyright 2015 Taylor & Francis

This is an Accepted Manuscript of an article published by Taylor & Francis in *Journal of Education Policy* on 2015, available online: <http://www.tandfonline.com/doi/full/10.1080/02680939.2015.1070206>

Notice: *Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source:*

<https://doi.org/10.1080/02680939.2015.1070206>

Who pays for standardised testing? A cost-benefit study of mandated testing in three Queensland secondary schools.

Merilyn Gladys Carter (Corresponding author)

Faculty of Education, Queensland University of Technology, Brisbane, Australia

merilyn.carter@qut.edu.au

Valentina Klenowski

Faculty of Education, Queensland University of Technology, Brisbane, Australia

val.klenowski@qut.edu.au

Christina Chalmers

Faculty of Education, Queensland University of Technology, Brisbane, Australia

c.chalmers@qut.edu.au

Disclosure Statement

The authors are not aware of any conflicts of interest, actual or potential, either personally or institutionally, that may exert an influence on the study upon which this article is based.

Who pays for standardised testing? A cost-benefit study of mandated testing in three Queensland secondary schools.

This paper reports on an Australian study that explored the costs and benefits of the National Assessment Program, Literacy and Numeracy (NAPLAN) testing, both tangible and intangible, of Year 9 students in three Queensland schools. The study commenced with a review of pertinent studies and other related material about standardised testing in Australia, the USA and UK. Information about NAPLAN testing and reporting, and the pedagogical impacts of standardised testing were identified, however little about administrative costs to schools was found. A social constructivist perspective and a multiple case study approach were used to explore the actions of school managers and teachers in three Brisbane secondary schools. The study found that the costs of NAPLAN testing to schools fell into two categories: preparation of students for the testing; and administration of the tests. Whilst many of the costs could not be quantified, they were substantial and varied according to the education sector in which the school operated. The benefits to schools of NAPLAN testing were found to be limited. The findings have implications for governments, curriculum authorities and schools, leading to the conclusion that, from a school perspective, the benefits of NAPLAN testing do not justify the costs.

Keywords: Standardised testing, mandated testing, schools, cost-benefits, NAPLAN, secondary schooling

Introduction

Mandated standardised testing is a feature of many educational systems. Whilst there has been considerable discussion of the impact of such testing on teaching and learning, there are few studies that explore the broader implications for schools. By examining the costs and benefits of mandated standardised testing in three secondary schools in Queensland, Australia, this paper considers whether the substantial costs to schools are justified by the benefits. Although it is acknowledged that the costs and benefits of mandated standardised testing affect a range of individuals and groups, the focus of this paper is the school

perspective. It draws on a doctoral study of changes in secondary schools associated with mandated standardised testing in Australia (Carter 2014).

The costs associated with standardised testing for schools can be classified in several categories. *Direct costs* are the most easily quantified as they are identifiable cash outlays, for example, the purchase of stationery for student use during the tests. Some resources, such as staff time, are shared between testing and other school activities, requiring the apportionment of costs. *Indirect costs* are linked to the standardised testing program, but not necessarily to the execution of the tests. The most significant indirect cost for most schools is the preparation of students for standardised testing. Indirect costs can be further classified as *tangible* and *intangible*. Tangible costs can be quantified, for example, the purchase or printing of practice tests, whereas a dollar value cannot be attributed to intangible costs such as the loss of teaching time in other areas if standardised testing activities take priority. Economists call them *opportunity costs*. Although unquantifiable, intangible costs are significant and must be included in an analysis of the cost of a standardised testing program for schools.

Whilst there have been many studies of the indirect costs of standardised testing in areas such as teaching and learning, there have been few studies that consider, at the school level, the direct costs and compare the full range of costs to the benefits.

The Queensland context

Concerns that Australian school outcomes were falling behind those of its neighbours in the Asian region, supported by the results of international testing (for example, OECD 2013), have led to the standardisation of school education throughout Australia. The impact of such comparative analyses include the development of an Australian Curriculum (2014a), and its progressive implementation since 2012, by a government agency called the Australian Curriculum Assessment and Reporting Authority (ACARA). Another feature of the national

approach was the adoption in 2008 of mandated standardised testing in every Australian school, known as the National Assessment Program, Literacy and Numeracy (NAPLAN), now also managed by ACARA. The introduction of NAPLAN testing provided, for the first time, nationally comparable data on individual student performances (ACARA 2014c).

The point-in-time tests are undertaken over three consecutive days in mid-May each year by students in Years 3, 5, 7 and 9 (QSA 2014). The tests assess the domains of reading, writing, grammar and punctuation, spelling, and numeracy. The results are reported in several ways: (a) detailed outcomes for every student are supplied to schools and parents; (b) a comprehensive national report focusing on aggregate outcomes for the eight states and territories is published annually (see ACARA 2014d); and (c) a summary of the results in each domain and year level of every Australian school are posted on a website administered by ACARA known as *My School* (ACARA 2014b). These results are regularly summarised in newspapers as simplistic league tables of schools (for example, Ferrari and Knapp 2014), in apparent breach of the conditions of entry to the My School website. Until 2014, some Australian Government funding was provided to state and territory educational systems and to disadvantaged and/or underperforming schools aimed at supporting the achievement of specified targets that include NAPLAN results (COAG 2008a, b, Lingard and Sellar 2013). The publication of individual schools' results and the linking of additional funds to test results for some schools makes NAPLAN high-stakes testing for teachers and schools, but not for students.

Following the embarrassment of poor performances by Queensland students in the first set of NAPLAN tests in 2008, which has continued in subsequent years, test score improvements became part of the Queensland Government's political agenda (Lingard and Sellar 2013, Hardy 2014). In the schools run by that government (approximately 70% of all Queensland schools), NAPLAN test results have been used to develop targets for the schools

and their students. Since 2010, all government schools have been subject to quadrennial teaching and learning audits that focus, in part, on improvements in measurable student outcomes (including NAPLAN test results) (DETE 2014b). More explicitly, government schools are expected to continually improve their NAPLAN results relative to the mean (DETE 2012), a problematic goal given the norm-referenced nature of the test results (Darling-Hammond 2007). They are required to develop plans to ensure that every student can meet national minimum standards in literacy and numeracy (DETE 2014a), measured by ACARA as the achievement of a minimum number of correct responses in the NAPLAN tests (Carter 2014). Consistent failure to meet these goals can lead to a review of the performance of the school principal (DETE 2012). Linking principals' future employment to an expectation of continual improvement in measurable data, including NAPLAN test results, with little apparent consideration of the research literature questioning the validity of such an approach (for example, Stobart 2008, Wu 2010, Boston 2009), can result in a disproportionate focus on those activities that can be measured¹.

ACARA is responsible for the development of NAPLAN tests, the compilation of the test data and national reporting of student results. Half of ACARA's funding is provided by the Australian Government, with the remainder shared between the states and territories on a proportionate basis. ACARA has reported that its annual costs in relation to NAPLAN testing are generally between 7 and 7.5 million Australian dollars (ACARA 2014c). As approximately 1.05 million students attempted the NAPLAN tests in 2013 (ACARA 2013a), that represents an approximate cost of \$A7 per student. The administration of the NAPLAN tests, including printing, distribution and marking of test papers, is undertaken and financed

¹ The Queensland Government changed following elections in January 2015. It is too soon to identify if that change of government has resulted in a significant change in the policies affecting government schools.

by the state and territory curriculum authorities². The Queensland Government in 2009 advised that their average cost per student of NAPLAN testing was approximately \$A20 (Queensland Parliament 2009, 1).

The costs of NAPLAN testing for governments would usually go beyond the scope of this paper which focuses on the costs for schools. However, in 2009 a NAPLAN administration charge was introduced by the Queensland Government to recover approximately sixty per cent of that government's costs of administering the NAPLAN tests. The annual charge ranges from \$A19 to \$A21 per student, depending on the costs incurred in any year. Whilst the charge nominally applies to all students enrolled in the NAPLAN testing years, it is not deducted from government school budgets as the government has decided not to recoup the costs from its own schools. However, the charge is recovered from the non-government schools (by invoice either directly to the school, in the case of independent schools, or to the governing body for the school system, in the case of Catholic schools).

Until 2014, selected schools received some funding from the Australian Government to assist in improving literacy and numeracy outcomes. However, these funds were provided for specific additional programs (for example, professional development of teachers) and could not be used to offset the NAPLAN administration charge. Further, most of that funding was directed at government schools, which were not required to pay the NAPLAN administration charge.

The NAPLAN administration charge has not been widely publicised. Extensive searches for information about the charge (for example, of the websites and annual reports of government agencies, government budget papers, parliamentary proceedings, legislation, and

² The curriculum authority in Queensland is the Queensland Curriculum and Assessment Authority (QCAA), known before 1 July 2014 as the Queensland Studies Authority (QSA).

the internet generally) failed to yield any pertinent information other than an answer to a 2009 Question on Notice in the Queensland Parliament. The factual information about the charge was obtained from a written response to specific questions addressed by the study to the Queensland Minister for Education (Nov 10, 2014).

Costs of Standardised Testing on Curriculum and Pedagogy

The costs of standardised testing for schools are not limited to the direct costs associated with the tests. Many studies have analysed the indirect costs of standardised testing in areas such as teaching and learning.

If standardised test results are used to judge and compare schools (as in Australia), it leads to a disproportionate focus on test results. It assumes that test performance in a limited number of areas measures the quality of education (Kohn 2000) and that academic achievement is the most important objective of education (Sireci 2005). It ignores the development of students in non-academic areas and fails to address other worthwhile goals of schooling (Luna and Turner 2001, Barrier-Ferreira 2008). With more students remaining at school for longer, many students now have educational objectives other than academic achievement (Boston 2009). The public release of standardised test results for each school affects the reputation of the school, teachers and/or students, potentially impacting on future school enrolments and future employment prospects for teachers and school managers.

Pressure on teachers and principals to improve test results can distort their behaviour (Nichols and Berliner 2007, O'Neill 2013, Stobart 2008). It can lead to the prioritising of improvements in test scores by any means (Kohn 2000, Nichols and Berliner 2007, Samuels 2011) rather than focusing on learning, often labelled as teaching to the test. Many teachers have responded to the dilemmas associated with high-stakes testing (Miyasaka 2000) by adapting aspects of their instruction to improve test scores rather than learning (Abrams and Madaus 2003, House of Commons (UK) Children Schools and Families Committee 2008,

Mertler 2011). It has resulted in a greater focus on didactic methods such as drills of test items (Firestone, Monfils, and Schorr 2004). Madaus (1988) observed that everywhere that high-stakes tests operate, a tradition of practising past test papers has developed.

Whilst some preparation of students for any form of testing is appropriate, “there is a fine line between teaching to the test and teaching the test” (Nichols and Berliner 2007, 122) that may be crossed when the time devoted to test preparation is excessive. Many studies have revealed that the amount of class time devoted to standardised test preparation can be substantial (Nichols and Berliner 2007, Dulfer, Polesel, and Rice 2012, Thompson and Harbaugh 2013, Wyn, Turnbull, and Grimshaw 2014). A study of standardised test preparation practices concluded that “for teachers the stakes are high, and they react by doing what is necessary to prepare children to take the external tests” (Smith 1991, 525). As performance in external tests was influenced by variables outside the direct control of teachers, such as students’ socioeconomic backgrounds, they tended to regard the test results as a worthless measure of student achievement (Smith 1991, 538). Consequently, teachers used highly focused methods to prepare students with the least impact on their ‘normal’ teaching (implying that they considered test preparation and teaching the curriculum to be different). These practices confirm Moore’s (1994) finding that many teachers adopted test preparation practices that test specialists considered inappropriate.

The reduction of resources (such as teachers, funding, and class time) devoted to school activities that are not subject to standardised testing (such as foreign languages, liberal arts, or physical education programs) appears to be widespread. It occurs particularly in primary schools where teachers have more flexibility in the allocation of class time (Boston 2009, Boyle and Bragg 2006, Nathan 2008, Nichols and Berliner 2007). The reduction in resources for other learning areas is likely to have had an impact on students’ knowledge and

skills in those areas, and also on their development as well-rounded individuals (Boston 2009).

The unintended consequences of NAPLAN testing in schools were considered by an inquiry into the effectiveness of NAPLAN by a Committee of the Australian Senate. It found that:

The committee are [sic] also concerned with the comments made by ACARA that suggest unintended consequences can be as a result of a miscomprehension over what NAPLAN's intended consequences are. ... To suggest that changes in the classroom are not as a result of NAPLAN is not taking full responsibility for the profound impact that standardised testing can have. This in itself is not a reason not to test, but it is something that educational authorities need to be cognisant of in providing support to schools as part of the NAPLAN process. (Australian Senate 2014, 14)

The review of literature identified the indirect costs of standardised testing on teaching and learning in schools. However, limited information was found about the other costs of mandated standardised testing to schools including the requirement to administer those tests. As some of these costs are determined by the policies of governments and/or education authorities, they can vary between countries, states and territories. This article considers the costs and benefits for schools of standardised testing, both direct and indirect, in Queensland, using three secondary schools as examples.

Methodology

A social constructivist perspective was used to examine the actions of school leaders and teachers in three Queensland secondary schools. As the aim of constructivist inquiry is to understand what occurs (Guba and Lincoln 1994), a qualitative approach involving a multiple case study (Simons 2009, Stake 1995) was used to explore the practices of three schools with different governance arrangements and diverse approaches to NAPLAN testing. Data collection methods included semi-structured interviews and exchange of emails with selected

school managers and teachers, observation of school activities, and documentary analysis. Whilst the focus of the study was on the NAPLAN numeracy tests, much of the information collected was also relevant to the literacy testing. Coding and categorisation allowed the identification of similarities and differences between the three schools and the development of general propositions about the costs to schools of NAPLAN testing (Miles and Huberman 1994).

The three cases were selected on the basis of maximum variation (Miles and Huberman 1994) to compare and contrast the practices in different types of schools in the context of Year 9³ NAPLAN testing and reporting in 2013. As the providers of school education in Australia have traditionally belonged to one of three sectors: government; Catholic; and independent; a school from each sector was selected. They were all large coeducational secondary schools in the outer fringes of Brisbane (the largest city in Queensland), operating in the same political and educational policy context. They all described themselves as serving their local communities and did not typically draw students from outside their local catchment areas.

The schools had different governance arrangements, generally reflecting the system to which they belonged. The government school was owned by the Queensland Government and was required to implement the education policies of that government. Christian churches owned the two non-government schools: one Catholic and one Protestant (although not every independent school is affiliated with a particular faith). The principals of the two non-government schools had more autonomy in the management of their schools, especially in curriculum and pedagogical issues. Key characteristics of the three schools at the time of the study are summarised in Table 1.

³ At the time of the study, secondary education in Queensland was considered to be Years 8 (aged 12 or 13) to 12 (aged 17 to 18).

Table 1: Key characteristics of the schools studied, 2013

| Characteristic | Government School | Catholic School | Independent School |
|-------------------------------------------|-------------------|-----------------|--------------------|
| Student population (nearest 10) | 1000 | 870 | 1770 |
| Socio-economic status | low | average | high |
| School income per student (\$A, 2012) | \$15630 | \$13300 | \$14200 |
| Government funding (% of school income) | 96% | 74% | 58% |
| Full-time equivalent teachers (nearest 5) | 90 | 65 | 125 |
| Teacher-student ratio | 1:11.1 | 1:13.7 | 1:14.1 |

The three schools had different socio-economic characteristics (ACARA 2014b), reflecting the nature of the communities that they served and the requirement to pay fees for attendance at the non-government schools. A review of the literature by Toutkoushian and Curtis (2005) showed convincingly that socioeconomic factors are correlated highly with student, and hence school, success in standardised testing. Their own study demonstrated that unemployment rates, adult education, and parent income accounted for more than half of the variation in average standardised test scores.

The three schools had different approaches to NAPLAN testing. The government school, with a diverse student population assessed as having low levels of literacy and numeracy skills, responded to government policies by focusing intensively on improving those skills. Consequently, the school used NAPLAN outcomes as indicators of success. The Catholic school, with a history of NAPLAN outcomes that the school managers considered to be satisfactory, did not afford a high priority to NAPLAN processes, although many mathematics lessons were devoted to NAPLAN preparation. The independent school adopted a low-key approach that sought to minimise the pressure on teachers and students associated with NAPLAN testing. However, when many teachers misinterpreted that approach as a *laissez-faire* policy, the school managers identified a need to clarify their

expectations by including explicit statements about the preparation of students for NAPLAN testing in school curriculum documents.

Costs of NAPLAN testing

The three schools did not keep their financial accounts in a form that allowed the costs associated with NAPLAN testing to be extracted easily. Expenditure on NAPLAN testing was absorbed into other financial allocations (for example books of NAPLAN numeracy practice tests used by the Catholic school were purchased using mathematics department funds). Some equipment purchased to support the NAPLAN testing program (such as teaching resources and educational technology) was used over several years and/or also applied to other purposes, requiring apportioning of the costs. Other costs, such as lost teaching time, were intangible. Accordingly, it was difficult to isolate and quantify many of the costs associated with NAPLAN testing in a particular year. In some cases the costs were estimated by the school managers making the expenditure decisions, in others it has only been possible to describe the nature of the cost.

The NAPLAN administration charge collected by the Queensland Government has already been discussed. As non-government schools were unable to opt out of NAPLAN testing, they were compelled to fund a service that they may not have wanted. In the Catholic school system, where the charge was paid centrally with a consequent reduction in the funds distributed to individual schools, school managers were not aware that their school paid for NAPLAN testing.

The other costs of NAPLAN testing for schools fell into two categories: preparation of students for the tests and administration of the tests. They are discussed in the following sub-sections.

Preparation of students for NAPLAN testing

The review of literature earlier in this paper showed that the preparation of many students for high-stakes standardised tests focussed on raising test scores in a way that detracted from the usual progression through the curriculum content and skills. In this form of test preparation, often called “bolt-on” preparation, the primary lesson objective is to prepare students for the tests, with teaching the curriculum a secondary consideration. It typically includes the decontextualised practice of tests in similar formats, either during lessons or in full cohort test rehearsals.

All three schools studied engaged in forms of bolt-on NAPLAN preparation. Some teachers of English and mathematics used NAPLAN practice tests in their lessons. Both the government and independent schools rehearsed all five NAPLAN tests using past test papers under the same test conditions as would apply authentically in the NAPLAN tests. The Catholic school rehearsed one of the five tests under test conditions, with the other four tests practised during lessons. However, in the twelve months leading up to the 2013 Year 9 NAPLAN tests, the Coordinator of Mathematics in that school scheduled several weeks of mathematics class time for bolt-on test preparation. One of the teachers from the Catholic school described a more general feeling about the impact of bolt-on preparation: “I just want to make the observation that [the block of NAPLAN preparation time] has cut into teaching time. What we would call our plan, what we are trying to get through for the year, it puts a significant dent in that.”

The study found that most teachers saw test preparation only in terms of bolt-on approaches, illustrated by this comment by a teacher in the independent school:

I have, basically, the bottom Year 9 class. I’m getting some success with them this year, but with a class of that ability it’s a struggle to keep up. So, we just simply haven’t had the time to be able to do [NAPLAN preparation]. I’m struggling on a week-to-week basis. I’m keeping up, but without any spare time. As far as the NAPLAN preparation is concerned, I allocated two days just in the week before the NAPLAN tests.

Two days and I put together a selection of past NAPLAN questions for the kids to do and I had some homework based on that, as well. So that was basically what I covered. But, as I said, ideally, it would be nice to have sufficient time to be able to do it on a more regular basis.

It suggests that professional development was needed in methods of test preparation that do not detract from the aim of developing curriculum content and skills.

The alternative to bolt-on preparation, often called “built in” test preparation, occurs when the test preparation occurs as part of the study of curriculum content and skills. All school managers interviewed supported built in NAPLAN preparation, exemplified by this comment by a principal:

We are working with Heads of Department and Middle School coordinators to try and get them to do the NAPLAN preparation as part and parcel of what we do as English or maths teachers. I'd hate kids to know that today we are doing NAPLAN or today we are doing this kind of preparation.

The study found that these expectations translated into action in only a small proportion of classrooms. Although some teachers argued that good pedagogy was the only preparation needed for NAPLAN testing: “every lesson is a NAPLAN lesson really, to some extent”, few teachers could identify anything specific that they did in their day-to-day teaching of the curriculum to prepare students for NAPLAN testing.

As bolt-on preparation focuses primarily on the NAPLAN tests, thus diverting resources from other teaching and learning activities, the cost of this form of preparation was viewed as an indirect cost of testing. However, built in preparation, where the primary lesson objective is to teach particular curriculum objectives, was not considered to be a cost of testing (Carter 2014). Bolt-on practice that occurred during lessons resulted in the opportunity cost of teaching time that could otherwise have been devoted to the curriculum. Tangible costs of bolt-on preparation included the purchase of books of practice tests for each student (in the Catholic school), the reproduction of past test papers (in the independent

school), and the purchase of educational technology such as audience response systems (in the government school). Test rehearsals demanded similar resources as the conduct of NAPLAN tests, discussed below. However, they also involved additional costs, including the photocopying of past test papers, marking of practice tests, and the review of the test results with students.

Administration of the tests

All schools in Australia are required to administer NAPLAN tests on behalf of ACARA, including the provision of all on-site administrative support for the tests. Each year ACARA issues detailed instructions for the conduct of the tests (for example, ACARA 2013b) that make demands on teachers' time and other school resources. School principals are responsible for the management and conduct of the NAPLAN tests in their schools, but in each of the three cases that responsibility was delegated to a senior teacher. The administration of the Year 9 NAPLAN tests resulted in significant costs for each of the schools. They can be analysed according to their timing: before; during; and after the tests.

Preparation work for the tests in schools included the familiarisation of school managers with the extensive ACARA instructions for the conduct of the tests, the receipt, unpacking and distribution of the test papers, planning and notification of invigilation arrangements and related changes to teachers' duties, planning and dissemination of changes to student routines for the three test days, preparation of the room(s) used for the tests (for example, set up and/or rearrangement of desks, removal of wall posters), planning the seating arrangements of students in the test room(s), and arranging special test conditions for eligible students.

During the tests, additional activities included supervision of the entire testing process to ensure teacher and student compliance with test procedures, the loan of essential equipment (stationery and calculators) to those students who did not have them, accounting

for students' attendance and absences, and dealing with any student behaviour management issues. Test invigilation was generally undertaken by the teachers who would otherwise have been teaching the Year 9 students, without additional cost to the school; however, the independent school employed three additional teacher aides to assist with test invigilation and processing test papers.

Post-test activities included the collection, checking, counting, repacking, and despatch of completed test papers, arranging for students absent on test days to catch up the missed tests, restoring the room(s) used for the tests, and (eventually) distributing individual student NAPLAN reports to parents. The independent school diverted grounds staff from their usual duties to set up, and later remove, desks and chairs in the sports centre used for the tests, in the other schools it was done by teachers and students just before the end of the school day preceding the tests (further reducing the time available for teaching and learning).

Except where indicated, school managers, teachers and teacher aides absorbed the NAPLAN administrative duties into their existing workloads. Whilst most of these activities did not add to the schools' expenditure, the increase in workloads resulted in an opportunity cost if other duties could not be undertaken or the additional workloads were met by sacrificing personal leisure time.

The schools incurred various direct costs associated with NAPLAN testing. In the lower-socio-economic government school, students were provided with breakfast before the tests and snacks and drinks between the tests, to improve their stamina and concentration. On completion of the last test, those students who had attended all tests were rewarded with pizza. In the Catholic school students celebrated the end of the tests with cake provided by the school. The independent school did not offer any inducements to students, but engaged additional teacher aides to assist with test invigilation and processing test papers. All three schools were required to purchase stocks of calculators and pencils for loan to those students

who failed to provide their own test equipment. Managers in all three schools estimated that their NAPLAN-related expenditure for additional staff, food, and equipment to be approximately \$A7 per student. Schools were not reimbursed for these outlays.

Intangible costs of NAPLAN testing were the suspension of Year 9 teaching activities for three mornings in the two non-government schools (where students returned to normal lessons in the afternoons), and for three full days in the government school (where there was an alternative program for Year 9 students in the afternoons of the test days). In two of the schools the tests were conducted in a single teaching space generally used for other purposes. The diversion of those spaces from their normal usage for three days was another opportunity cost. NAPLAN testing also influenced the programming of other school activities for Year 9 and in some cases, the entire school. Most teachers avoided scheduling Year 9 class tests or assignments close to the dates of the NAPLAN tests. Similarly, excursions, major sporting events, and whole-school activities could not be planned on NAPLAN test days.

Summary

Given that NAPLAN administration charge applied to some schools but not others, the direct costs of NAPLAN testing for schools varied according to the education sector in which the school operated. This study found that the direct costs were approximately \$7 per student tested in government schools and \$28 per student tested for non-government schools.

Additionally, there were many indirect costs for schools, summarised in Table 2.

Table 2: NAPLAN processes leading to additional intangible costs for schools

| Timing | Process |
|-----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Preparation of students for tests | <ul style="list-style-type: none"> • Professional development for teachers in methods of test preparation that do not detract from the need to teach curriculum content and skills • Bolt-on test preparation: practice tests and full cohort test rehearsals • Marking of practice tests • Review of practice test results by teachers and students |
| Arrangements prior to the tests | <ul style="list-style-type: none"> • Familiarisation of school manager(s) with test instructions • Receipt, unpacking, distribution of test papers • Seating arrangements for students in the examination centre • Planning and notification of invigilation arrangements and changes to teachers' duties • Planning and notification of variations to student routines for the three days • Preparation of the room(s) used for the tests • Arranging special test conditions for eligible students • Arranging food rewards for students attending on test days |
| Activities during the tests | <ul style="list-style-type: none"> • Read test instructions to students • Ensure compliance with all test instructions • Provide required stationery and calculators to students • Test invigilation • Account for students' attendance and absences • Manage any student behaviour management issues |
| Activities after the tests | <ul style="list-style-type: none"> • Collect, check, account for, pack, and despatch completed test papers • Follow up absent students to ensure that they catch up missed tests • Restore room(s) used for the tests • Distribute student NAPLAN reports to their parents |
| Other | <ul style="list-style-type: none"> • Suspension of learning opportunities for test participants during the three day test period • Diversion of examination room(s) from normal uses • Impact on programming of other school activities |

The unquantifiable nature of many of these costs resulted in a considerable understatement of the true costs of NAPLAN testing for schools, suggesting that the full cost of NAPLAN testing for schools was substantial. In contrast, the next section will show that the schools derived limited benefits from NAPLAN testing and the resulting data.

Benefits of NAPLAN testing

The benefits of NAPLAN testing to schools are threefold. First, all three schools used NAPLAN data to rate themselves against national benchmarks and comparable schools. For example, the government school promoted their favourable 2012 NAPLAN outcomes compared to other local schools on a noticeboard outside the school.

Second, it has been argued that standardised testing leads to improvements in schools' pedagogy and assessment, especially in literacy and numeracy (Stobart 2008, Nichols and Berliner 2007, Nathan 2002). However, the government school made literacy and numeracy a priority independently of NAPLAN testing. The test results were not needed to inform them of their students' problems or to provide the incentive to act. In all three schools, the school managers stated repeatedly that changes in their curriculum and pedagogies were not driven by a wish to improve NAPLAN outcomes.

Finally, the use of NAPLAN data for diagnosing student weaknesses was held by the Australian Government to be one of the major benefits for schools of NAPLAN testing. However, the diagnostic use of NAPLAN data appeared to be of limited benefit to any of the schools in this study. The four month delay⁴ in the provision of NAPLAN results for each student meant that they were out of date by the time they were received by schools. All three schools paid for additional standardised testing of their students, suggesting that NAPLAN data were inadequate for their needs.

The schools may have derived some benefit from the availability of NAPLAN data. However, it did not appear that these limited benefits justified the significant costs to the schools.

⁴ In Queensland in 2014 that delay was reduced to three months.

Discussion

NAPLAN testing was adopted by governments as part of a policy of standardising approaches to school education throughout Australia, in response to perceptions of declining standards in Australian school education. As in many other countries, mandated NAPLAN testing was implemented by politicians for political purposes, but was not generally sought or welcomed by educators (Comber and Nixon 2009). Consistent with the overseas experience, the public release of NAPLAN data has led to its use for a variety of purposes, some appropriate and others not (O'Neill 2013). These political issues have impacted on the activities of schools. Whilst some school managers sought to minimise the impact of NAPLAN testing on their teachers and students, in all three schools in the study the curriculum and/or pedagogies were adapted in the context of NAPLAN testing. The pressure, direct or indirect, on all schools, but particularly government schools, to improve their NAPLAN outcomes has resulted in increased time and money devoted to preparation for those tests.

Other costs to schools, both tangible and intangible, flowed from the requirement that all schools administer the NAPLAN tests to their students, acting as agents of ACARA and QCAA/QSA. The recovery of some NAPLAN costs by the Queensland Government required non-government schools to pay for a testing program that they neither sought nor controlled. This is another example of political decision making impacting on the costs of NAPLAN testing for some schools.

NAPLAN testing diverts resources from other school priorities. The high stakes and mandatory nature of NAPLAN tests has ensured that all schools must allocate some of their resources to administer the tests and/or achieve higher test scores. It is impossible to identify the specific impact of this diversion of resources on individual schools and their students,

other than to comment that the principal and teachers of a school may be in a better position to determine what is best for their students than governments.

This study has implications for existing research, governments, curriculum authorities and schools. Each of these are discussed in the following sub-sections.

Existing research

The studies that discussed the negative consequences of politically mandated standardised testing for schools (for example, Boston 2009, Kohn 2000, Nichols and Berliner 2007, O'Neill 2013, Stobart 2008) were generally confirmed by this study. NAPLAN test scores have been used to hold schools, and indirectly their teachers, to account. This is particularly the case in Queensland where the government demanded improvements in NAPLAN scores in all government schools. The use of narrow, unrealistic and possibly invalid targets, based largely on NAPLAN scores, has distorted the behaviour of some schools. There is evidence of simplistic remedies such as extensive bolt-on test preparation and narrowed educational objectives. Whilst there was no evidence in the three schools studied of the strategic exclusion of students from NAPLAN tests, or of cheating in those tests, it is not surprising that a small number of Australian schools have succumbed to these pressures (ACARA 2012).

Consistent with the overseas experience, NAPLAN testing has facilitated the identification of schools requiring additional support in literacy and numeracy. The government school in this study welcomed the additional funding that enabled a focus on improvements in literacy and numeracy. However, the school managers and teachers did not need NAPLAN testing outcomes to identify the need for improvement.

Governments

The cost of NAPLAN testing has several implications for governments. Consistent with the overseas experience, the public release of NAPLAN data on the My School website has led to its use for a variety of purposes, many inappropriate. Whilst all of the schools in the study accepted the need to be accountable for their overall performances, undue focus on their literacy and numeracy outcomes overshadowed other achievements that they considered to be equally important. This led to comparisons between schools that were inappropriate at best, and misleading at worst.

The My School website (ACARA 2014b) provides some factual information about each school so that the school context can be taken into account when comparing NAPLAN outcomes. However, there are limitations to that approach. For example, school priorities, selective enrolment policies, teacher experience and/or competence, student behavioural issues, parental and cultural influences all affect test outcomes but are not reflected in information provided in the My School website. Additionally, individual school financial information shown in the website cannot be directly compared because of the recovery of some Queensland Government NAPLAN costs from the budgets of some schools but not from others. The annual publication of simplistic league tables of schools in the press, in apparent breach of the conditions of entry to the My School website, further encourages inappropriate comparisons between schools.

Curriculum authorities

NAPLAN testing was imposed without consultation with schools or their representatives and without any formal processes for schools and teachers to provide feedback to the curriculum authorities responsible for the tests. As schools are required to invest considerable resources in the NAPLAN testing process, they (or their representatives) should be given the opportunity to contribute to the development of NAPLAN testing policy and to suggest

changes to make it a more useful tool for schools. The creation of a forum for this purpose could result in useful exchanges between schools, teachers and curriculum authorities and lead to improvements to the NAPLAN testing processes.

The limited benefits of NAPLAN test data for schools and teachers is an example of an issue that could be raised in such a forum. The fact that every school in this study used commercial tests to obtain additional information about their students is a reflection of the usefulness of NAPLAN data for those schools. The data would be more useful to schools if continued reductions in the delay between testing and the return of NAPLAN test data can be achieved⁵.

Schools

Schools, especially government schools in Queensland, have been subject to pressure to improve NAPLAN scores. The policy that all government schools in Queensland can continuously improve their NAPLAN outcomes relative to the mean is both unrealistic and suggests fundamental misunderstandings of the definition of a mean and the normalised scoring of NAPLAN tests. It fails to recognise that short-term set-backs are inevitable as academically stronger and weaker student cohorts progress through schools. Given the pressure to improve NAPLAN test scores, the challenge for principals is to resist the temptation to adopt ‘quick-fix’ remedies, such as extended periods of bolt-on preparation that exacerbate the unintended consequences of standardised testing, at the cost of learning the curriculum content and skills. Whilst they may seek to satisfy the short-term political imperative of immediate improvement in test scores, such remedies are not in the best interests of the overall development of students.

⁵ Options for online NAPLAN testing as a way of reducing the delay in returning test results to schools are currently being investigated by ACARA.

Every school is required to meet the cost of NAPLAN preparation and testing in their school. However, the imposition by the Queensland Government of a poorly publicised NAPLAN administration charge, on non-government schools only, is inequitable for three reasons. First, as noted above, it makes comparisons of individual school finances impossible if the cost of NAPLAN testing is recovered from the budgets of some schools but not others. It is argued that, if the Queensland Government charged its own schools for the costs of NAPLAN testing, it would be reimbursing itself. However, there are many precedents elsewhere for financial transfers between different branches of government⁶, justified in terms of transparency and accountability. Second, the administration charge transfers some of the cost of NAPLAN testing to the parents of students at non-government schools through their payment of school fees. Few parents would be aware that they are contributing to this cost. Finally, the requirement that non-government schools pay the majority of the Queensland Government's cost of testing of their students assumes, incorrectly, that schools are the major beneficiaries of NAPLAN information. The charge, with its lack of transparency and accountability, should be reviewed.

Conclusion

In this article we have discussed and analysed the tangible and intangible costs to schools associated with NAPLAN testing, which have been evaluated as substantial. The costs of NAPLAN have received little attention in the wider community. Given the compulsory nature and high cost of schools' participation in the NAPLAN testing program, it is disappointing that policy makers have failed to give more weight to the concerns of schools and teachers about that testing and the usefulness of the resulting data.

The study of only three schools in Queensland may limit the generalisation of the conclusions to other schools or jurisdictions, especially outside Australia. However, the

⁶ For example, the Australian Government pays itself taxes arising from its own business activities.

experience of these schools may resonate with the circumstances of other schools, or lead others to consider the investigation of similar issues in other contexts. It is for the reader to determine the applicability of this study beyond the three Queensland schools.

Standardised testing may provide useful information for those outside schools: governments; curriculum authorities; parents; researchers; the press and media; businesses; and the community. However, for schools, the benefits of that testing do not justify the substantial associated costs that they incur. There is a need for governments to reduce the cost burden of standardised testing on schools so that they can focus their limited resources on their core business – developing the citizens of tomorrow.

References

- Abrams, Lisa M., and George F. Madaus. 2003. "The lessons of high-stakes testing." *Educational Leadership* 61 (3):31-35.
- ACARA. 2012. Report of 2012 NAPLAN test incidents. Sydney, NSW: Australian Government.
- ACARA. 2013a. National assessment program literacy and numeracy: Achievements in reading, writing, language conventions and numeracy 2012. Sydney, NSW: Australian Government.
- ACARA. 2013b. National Assessment Program Literacy and Numeracy: National protocols for test administration 2013. Sydney, NSW.
- ACARA. 2014a. "The Australian Curriculum v 6.0." Education Services Australia, Last Modified February 18 Accessed June 14. <http://www.australiancurriculum.edu.au>.
- ACARA. 2014b. "My School." Accessed March 7. <http://www.myschool.edu.au/>.
- ACARA. 2014c. "NAP: National Assessment Program." ACARA Accessed June 14. <http://www.nap.edu.au/>.
- ACARA. 2014d. National assessment program literacy and numeracy: Achievements in reading, writing, language conventions and numeracy 2013. Sydney, NSW: Australian Government.
- Australian Bureau of Statistics. 2014. "Schools, Australia 2013." Australian Government, Last Modified Feb 6 Accessed June 14. <http://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/9448F2F814FA0311CA2579C700118E2D?opendocument>.
- Australian Senate. 2014. "Effectiveness of the National Assessment Program – Literacy and Numeracy." Parliament of Australia Accessed 4 Apr 2014. http://www.aph.gov.au/Senate/committee/eet_ctte/naplan/submissions.htm.
- Barrier-Ferreira, Julia. 2008. "Producing commodities or educating children? Nurturing the personal growth of students in the face of standardized testing." *Clearing House* 81 (3):138-140.
- Boston, Ken. 2009. "League tables." *Teacher: The National Education Magazine* 205 (Oct 2009):54.

- Boyle, Bill, and Joanna Bragg. 2006. "A curriculum without foundation." *British Educational Research Journal* 32 (4):569–582.
- Carter, Merilyn. 2014. "A multiple case study of NAPLAN numeracy testing of Year 9 students in three Queensland secondary schools." (Doctor of Philosophy, monograph), Education, Queensland University of Technology.
- COAG. 2008a. National partnership agreement on literacy and numeracy. Canberra ACT: Ministerial Council for Federal Financial Relations.
- COAG. 2008b. National partnership agreement on low socio-economic status school communities. Canberra ACT: Ministerial Council for Federal Financial Relations.
- Comber, Barbara, and Helen Nixon. 2009. "Teachers' work and pedagogy in an era of accountability." *Discourse: Studies in the Cultural Politics of Education* 30 (3):333-345. doi: 10.1080/01596300903037069.
- Darling-Hammond, Linda. 2007. "Race, inequality and educational accountability: the irony of 'No Child Left Behind'." *Race, Ethnicity & Education* 10 (3):245-260. doi: 10.1080/13613320701503207.
- DETE. 2012. United in our pursuit of excellence: agenda for improvement 2012-2016. edited by Department of Education Training and Employment. Brisbane, Qld: Queensland Government.
- DETE. 2014. "Great results guarantee." Queensland Government Accessed July 24. <http://education.qld.gov.au/schools/grants/state/targeted/great-results.html>.
- DETE. 2014. "Teaching and learning audits." Queensland Government Accessed March 30 2015. <http://education.qld.gov.au/nationalpartnerships/teach-learn-audits.html>
- Dulfer, Nicky, John Polesel, and Suzanne Rice. 2012. The impacts of high stakes testing on schools, students and their families: An educator's perspective. In *The experience of education*. Sydney NSW: Whitlam Institute, University of Western Sydney.
- Ferrari, Justine, and Peter Knapp. 2014. "Your School." *Weekend Australian*, June 21, 1-15. Accessed June 21.
- Firestone, William A, Lora Monfils, and Roberta Y Schorr. 2004. "Test preparation in New Jersey: Inquiry-oriented and didactic responses." *Assessment in Education: Principles, Policy & Practice* 11 (1):67-88.
- Guba, Egon G, and Yvonna S Lincoln. 1994. "Competing paradigms in qualitative research." In *Handbook of qualitative research.*, edited by Norman K Denzin and Yvonna S Lincoln, 105-117. Thousand Oaks, CA: Sage Publications, Inc.
- Hardy, Ian. 2014. "A logic of appropriation: enacting national testing (NAPLAN) in Australia." *Journal of Education Policy* 29 (1):1-18. doi: 10.1080/02680939.2013.782425.
- House of Commons (UK) Children Schools and Families Committee. 2008. Testing and assessment. edited by House of Commons. London UK: The Stationery Office Ltd.
- Independent Schools Queensland. 2013. "Title." Memorandum, Brisbane Qld.
- Kohn, Alfie. 2000. "Burnt at the high stakes." *Journal of Teacher Education* 51 (4):315-327. doi: 10.1177/0022487100051004007.
- Lingard, Bob, and Sam Sellar. 2013. "'Catalyst data': Perverse systemic effects of audit and accountability in Australian schooling." *Journal of Education Policy* 28 (5):634-656.
- Luna, Catherine, and Cara Livingstone Turner. 2001. "The impact of the MCAS: Teachers talk about high-stakes testing." *The English Journal* 91 (1):79.
- Madaus, G F. 1988. "The influence of testing on the curriculum." In *Critical issues in curriculum: eighty-seventh yearbook of the National Society for the Study of Education*, edited by Laurel N Tanner. Chicago, IL: University of Chicago Press.
- Mertler, Craig A. 2011. "Teachers' perceptions of the influence of No Child Left Behind on classroom practices." *Current Issues in Education* 13 (3):2-34.

- Miles, Matthew B, and A M Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*. 2nd ed. Thousand Oaks CA: Sage Publications Inc.
- Miyasaka, J. R. 2000. "A framework for evaluating the validity of test preparation practices." annual meeting of the American Educational Research Association, New Orleans.
- Moore, William P. 1994. "Appropriate test preparation: Can we reach a consensus?" *Educational Assessment 2* (1):51-68.
- Nathan, Linda. 2002. "The human face of the high-stakes testing story." *Phi Delta Kappan 83* (8):595-600.
- Nathan, Linda. 2008. "What's been lost in the bubbles." *Educational Leadership 66* (2):52-55.
- Nichols, Sharon Lynn, and David C Berliner. 2007. *Collateral damage: How high-stakes testing corrupts America's schools*. Cambridge MA: Harvard Education Books.
- O'Neill, Onora. 2013. "Intelligent accountability in education." *Oxford Review Of Education 39* (1):4-16. doi: 10.1080/03054985.2013.764761.
- OECD. 2013. PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know. Paris: OECD.
- QSA. 2013. Annual report 2012–13. Brisbane, Qld: Queensland Government.
- QSA. 2014. National assessment program literacy and numeracy: Handbook for principals 2014. Brisbane, Qld: Queensland Government.
- Queensland Parliament. 2009. Question on Notice No. 1725. Dec 11 2009. Retrieved from <https://www.parliament.qld.gov.au/documents/tableOffice/questionsAnswers/2009/1725-2009.pdf>
- Ryan, Steve. 2009. "EQ bungling causes NAPLAN fiasco." *Queensland Teachers' Journal*, March, 7.
- Samuels, Christina A. 2011. "Test-tampering found rampant in Atlanta system." *Education Week 30* (36):1.
- Simons, Helen. 2009. *Case study research in practice*. Los Angeles CA: SAGE.
- Sireci, Stephen G. 2005. "The most frequently unasked questions about testing." In *Defending standardised testing*, edited by Richard P Phelps, 111-122. Mahwah, NJ: Lawrence Erlbaum Associates.
- Smith, Mary Lee. 1991. "Meanings of test preparation." *American Educational Research Journal 28* (3):521-542.
- Stake, Robert E. 1995. *The art of case study research*. Thousand Oaks CA: Sage Publications Inc.
- Stobart, Gordon. 2008. *Testing times: The uses and abuses of assessment*. Abingdon UK: Routledge.
- Thompson, Greg, and Allen G. Harbaugh. 2013. "A preliminary analysis of teacher perceptions of the effects of NAPLAN on pedagogy and curriculum." *The Australian Educational Researcher 40* (3):299-314. doi: 10.1007/s13384-013-0093-0.
- Toutkoushian, Robert K., & Curtis, Taylor. 2005. "Effects of socioeconomic factors on public high school outcomes and rankings." *The Journal of Educational Research, 98*(5):259-271.
- Wu, M. 2010. "Measurement, sampling, and equating errors in large-scale assessments." *Educational Measurement: Issues and Practice, 29*(4):15-27.
- Wyn, Johanna, Malcolm Turnbull, and Lyndall Grimshaw. 2014. *The Experience of Education: The impacts of high stakes testing on school students and their families. A qualitative study*. Sydney, NSW: Whitlam Institute.

