Publication of School Results and League Tables Harm Education

Trevor Cobbold

SAVE OUR SCHOOLS

Research Paper

June 2010

http://www.saveourschools.com.au

Table of Contents

Key Points

1. Far from improving education, publication of school results and league tables are likely to significantly harm education. Overseas experience shows that they:
   - Narrow the curriculum;
   - Distort teaching practice;
   - Disadvantage low and high achieving students;
   - Unfairly stigmatise low achieving students;
   - Make it more difficult for low performing schools to retain high quality teachers;
   - Discourage co-operation and collaboration between schools and teachers; and
   - Increase social segregation and inequity in education.

2. The curriculum is narrowed in at least two ways:
   - More time is devoted to the tested subjects of literacy and maths while untested subjects such as science, history, social studies, languages, arts and music, physical education and health receive much less time. Even recess gets cut;
   - In the subjects that are tested, greater emphasis is given to the areas that are most conducive to testing by multiple-choice questions and there is less teaching of more complex skills.

3. Teaching is distorted because schools and teachers tend to respond to pressure created by publication of school results and league tables by focusing more on teaching test taking skills and practicing for tests.

4. Publishing school results and league tables can undermine efforts to improve the quality of teaching because they turn the notion of a "good teacher" into one who increases test scores.

5. Schools tend to concentrate on improving the results of students who are on the border of accepted benchmarks at the expense of both high and low achieving students.

6. Publication of the results of individual schools and public rankings of schools may unfairly stigmatise and humiliate some students and alienate them from schooling.

7. Published school results and league tables may be used as a job guide and make it more difficult for schools with low results to retain and attract quality teachers and principals.

8. Publication of school results and competition for league table rankings can reduce collaboration between schools and between teachers within schools.

9. Publication of school results and league tables tends to increase socio-economic and ethnic segregation between schools which exacerbates inequity in education because:
   - Student learning needs in some schools increase without proportionate increases in resources to meet those needs;
   - Increasing concentrations of students from low socio-economic status families in some schools tend to lead to lower overall outcomes.

10. There is evidence of some of these effects already in Australia after only one year of publishing school results.

Summary

Far from improving education, publication of school results and league tables are likely to significantly harm education. Overseas experience shows that they:

- Narrow the curriculum;
- Distort teaching practice;
- Disadvantage low and high achieving students;
- Unfairly stigmatise low achieving students;
- Make it more difficult for low performing schools to retain high quality teachers;
- Discourage co-operation and collaboration between schools and teachers; and
- Increase social segregation and inequity in education.

There is evidence of some of these effects already in Australia after only one year of publishing school results.

Narrowing of the curriculum

Publication of school results and league tables restrict student learning because they narrow the curriculum and teaching. Students receive a less rounded education.

There are two major aspects of narrowing the curriculum. First, more time and resources are devoted to the tested subjects while other subjects not tested are neglected. Overseas evidence shows that schools direct more resources into the tested subjects of literacy and maths while untested subjects such as science, history, social studies, languages, arts and music, physical education and health receive much less time. Even recess gets cut.

Second, within those subjects that are tested, there is greater focus on the skills that are most conducive to testing by multiple-choice questions and less teaching of more complex thinking and writing skills.

There is already some evidence that publishing NAPLAN school results on My School is narrowing the curriculum in secondary schools in Australia. A survey conducted by the Australian Secondary Principals Association found that 33% of principals said that publication of school results had reduced the breadth of curriculum in their school.

Distorting teaching practices

Teaching practice also tends to change under the pressure of publishing school results and the pressure to improve league table rankings.

There is extensive evidence from overseas that schools and teachers tend to respond to publication of school results and league tables by focusing more on teaching test taking skills and practicing for tests at the expense of deeper learning experiences that develop analytical skills and greater understanding. League tables turn classrooms into test preparation factories. Weeks and months are devoted to test preparation.

There is considerable evidence that publication of NAPLAN school results on the My School website and publication of school league tables is causing schools to spend large amounts of time on test practice in class to the detriment of other areas of the curriculum. A survey conducted by the Australian Secondary Principals Association found that 65% of schools reported that they had increased the time spent in class on preparation for the NAPLAN tests in 2010 and 70% said they had increased the time spent on practising tests. A report

published by the Australian Primary Principals Association said large amounts of valuable instructional time was taken up by coaching and practising tests.

Many teachers now refer to NAPLAN as “napalm” because “it kills everything in the classroom”.

There is also evidence that school administrators in some states have pressured principals and schools to devote more time to practising for tests.

This focus on test preparation and raising test scores may undermine efforts to improve the quality of teaching in schools. Some studies show that the idea of a “good teacher” changes in schools under the pressure to lift school results. A “good teacher” becomes someone who increases test scores rather than someone who guides students to deeper and wider learning experiences and encourages self-motivated learning.

Disadvantages low and high achieving students

The pressure to improve school results can also create incentives for schools to ignore students at the lower and upper ends of the student achievement scales. Several English and US studies have found that schools concentrate on improving the results of students who are on the border of accepted benchmarks and neglect the lowest and highest achievers.

Improving the results of students just below benchmarks is seen as the most efficient way to increase a school’s average score or the proportion of students achieving a benchmark. However, while this strategy gives schools a better chance of improving their ranking, it may lead to worse outcomes for low and high performing students.

In addition, publishing outcomes of individual schools and public rankings of schools often unfairly stigmatises and humiliates schools, teachers, students and their families. Students and teachers in particular years may be easily identifiable as the “culprits” when a school gets a low ranking, especially in small schools, of which there are many in rural areas of Australia. Students who are humiliated for their lower learning accomplishments are unlikely to respond positively in their future learning.

Harder for low performing schools to retain quality teachers

League tables of school results serve as a job guide for teachers to apply to highly ranked schools with fewer learning and behavioural problems. This means that low ranked schools often end up with the least qualified, least experienced teachers.

Discourages collaboration and co-operation

Collaboration is generally seen as an important way to spread innovative approaches and good teaching practice both between and within schools. Publication of school results and competition for rankings can reduce collaboration between schools and teachers and slow the dissemination of best practice.

Schools will be reluctant to share successful practices with other schools if it means those schools could leapfrog them on league tables. A lower ranking for a school could result in a decline in enrolments, less financial resources as students leave and greater difficulty in holding and attracting staff.

There is also evidence from overseas that focus on school results undermines co-operation and collaboration between teachers within schools. The pressure to improve school results

can generate a climate of fear amongst teachers, undermine trust and result in a breakdown of the professional and social relationships needed to sustain collective professional support for student development.

Increases social segregation and inequity in education

Publication of school results and league tables tends to increase social segregation in and inequity in schooling. This is driven by both schools and parents.

Publishing school results and league tables create greater incentives for some schools to choose their students to maximise their results. There is abundant evidence from overseas of schools of “cream skimming” students most likely to achieve good results – these students are generally from higher income families.

Many parents see education as a “positional good” in that the value of education depends not on the learning acquired but on the relative status of the school attended. Publishing tables of school results aids this search for status and self-segregation. In general, it is well-off families who make greatest use of choice of school and finding the "best" schools inevitably means those that have less low-income and minority students.

Studies show that students from higher income families are far more likely to transfer to wealthier school districts and that white students are more likely to opt out of racially diverse schools and transfer to those with greater percentages of white students.

Increased social segregation in schooling induced by choice and competition between schools and aided by publication of school results and league tables can exacerbate inequity in education in two main ways.

First, it increases disparities between schools in student learning needs and the real resources available to meet those needs. Low SES schools are generally funded on the same per capita basis as other schools, with few allowances for the level of need they have to deal with. They have less real resources because they have higher costs and burdens. They also tend to have less experienced, less qualified teachers.

Second, increasing concentrations of students from low socio-economic status (SES) families in some schools tend to lead to lower overall outcomes. Schools with high concentrations of socio-economically and educationally disadvantaged students often have detrimental effects on student achievement. A student attending such a school is likely to have lower outcomes than a student from a similar background attending a school where the average SES of the student body is high.

This impact also appears to be greater for low SES and immigrant and minority students. There is a "double jeopardy" effect for these students in that they tend to be disadvantaged because of their circumstances at home, but when they are also segregated into low SES and/or predominantly minority schools this disadvantage is compounded.

A further effect of increased segregation in schools by class, religion and race is to make it more difficult for children to develop a real understanding of people of different backgrounds and to break down barriers of social intolerance. Increased social segregation means that more and more students have less experience of mixing with and learning with children from different socio-economic, cultural and ethnic backgrounds.

1. Introduction

The weight of evidence from the major research studies over the past 20 years is that publication of school results and league tables do not increase student achievement. But, far from improving school results, there is also strong evidence that they do significant harm to education. Overseas experience shows that they:

- Narrow the curriculum;
- Distort teaching practice;
- Disadvantage low and high achieving students and unfairly stigmatise low achieving students;
- Make it more difficult for low performing schools to retain high quality teachers;
- Discourage co-operation and collaboration between schools and teachers; and
- Increase social segregation and inequity.

There is some evidence of this already happening in Australia after only one year of publishing school results.

2. Narrowing the curriculum

Publishing school results creates incentives for schools to narrow the curriculum and provide a less rounded education. There are three major aspects of narrowing the curriculum:

- More time and resources is devoted to the tested subjects while other subjects not tested are neglected;
- In the subjects that are tested, greater emphasis is given to the areas that are most conducive to testing by multiple-choice questions and there is less teaching of more complex thinking and writing skills

Overseas evidence shows that schools direct more time into the tested subjects of literacy and maths while untested subjects such as science, history, social studies, languages, arts and music, physical education and health receive much less time. Even recess gets cut.

A major review of the primary school curriculum in England found that testing and publishing school results in English and maths has distorted children's learning and eroded their entitlement to a broad education [Alexander 2009; see also Alexander 2010]. It said that children were receiving an education that was "fundamentally deficient". It was neither broad nor balanced, and it valued memorization and recall over understanding and inquiry. The review found upper primary school students spend around half their time in the classroom studying English and maths because of the pressure to pass national tests. As a result, other subjects such as history, geography, art and science have been "squeezed out" of the curriculum.

> ....as children move through the primary phase, their statutory entitlement to a broad and balanced education is increasingly but needlessly compromised by a 'standards' agenda which combines high-stakes testing and the national strategies' exclusive focus on literacy and numeracy.
> The most conspicuous casualties are the arts, the humanities and the kinds of learning in all subjects which require time for talking, problem-solving and the extended exploration of ideas. [CPR: 22]

The leader of the review, Professor Robin Alexander, has said that the narrowing of the curriculum may have even reduced overall standards:

> The Cambridge Review's evidence shows how the pursuit of this narrow concept of 'standards' at the primary stage, in which test scores in literacy have been treated as proxies for the quality of primary

> education as a whole, has over the past 13 years seriously compromised children's legal entitlement to a broad and balanced curriculum. We also consider it possible that because standards in the basics and the availability of a broad and balanced curriculum have been shown empirically to be linked, the narrowing of the curriculum in pursuit of standards in 'the basics' may have had the opposite result to that intended, *depressing* standards in 'the basics' rather than raising them. As collateral damage goes, that's pretty spectacular. [Alexander 2010: 6]

The study confirmed the findings of other academic studies. For example, an earlier study of longitudinal data on the national curriculum from 1997 to 2004 found a primary curriculum skewed in the direction of English and mathematics to the detriment of science, the humanities and the arts [Boyle & Bragg 2006]. English and mathematics had increased in percentage teaching time while science, geography and history had decreased. It said this was caused by a range of central policy requirements including national testing and publishing its results.

The UK Qualifications and Curriculum Agency (QCA) submitted to a House of Commons inquiry on testing and assessment that 90% of primary and 79% of secondary schools had reported to the QCA that national testing has led to students being offered a narrower curriculum [HC 2008b: Ev 22]. The former head of the QCA has said:

> In England, the government's use of the key stage tests has seriously damaged the breadth and quality of primary education....As a result the school curriculum is narrower and poorer than it was when the tests were introduced in 1997. In many schools, the time spent on areas of the curriculum which are not externally assessed has contracted sharply. [Boston 2009: 4-5]

The Commons inquiry report concluded:

> A creative, linked curriculum which addresses the interests, needs and talents of all pupils is the casualty of the narrow focus of teaching which we have identified. Narrowing of the curriculum is problematic in two ways: core subjects are emphasised to the detriment of other, important elements of the broader curriculum; and, for those subjects which are tested in public examinations, the scope and creativity of what is taught is compromised by a focus on the requirements of the test. We are concerned that any efforts the Government makes to introduce more breadth into the school curriculum are likely to be undermined by the enduring imperative for schools, created by the accountability measures, to ensure that their pupils perform well in national tests. [HC 2008a: para 140]

There is similar evidence of an increasingly unbalanced curriculum in the United States as a result of the pressure to lift test results in English and mathematics.

A study prepared for the U. S. Department of Education as part of its requirement under the No Child Left Behind legislation to report to Congress on the effects of the law found substantial evidence that time for non-tested subjects, such as art, social studies and foreign languages was sacrificed in favour of tested subjects [Hannaway & Hamilton 2008].

> Not surprisingly, the available evidence confirms that, under conditions of a performance-based accountability system, instructional time on tested subjects increases. It also appears that this increase sometimes comes at the expense of non-tested subjects, resulting in a narrowing of the curriculum. [7]

A report by the National Centre for Education Statistics found that instructional time in grades 1 through to 4 on English and mathematics increased between 1987-88 and 2004-04 but declined in science and social science [Morton & Dalton 2007].

A study by the US Centre on Education Policy has showed that since the No Child Left Behind Act was enacted in 2002 to require more testing and reporting for reading and mathematics, by 2007 58% of all school districts had increased the time spent in primary

schools on English and 45% had increased the time spent on mathematics [McMurrer 2008]. The average increase across these districts amounted to 141 extra minutes per week (or an average of 28 minutes per day) in English and 89 extra minutes per week (or about 18 minutes per day) in maths. These amounted to increases of 47 and 37% respectively.

The time spent on social studies was reduced in 36% of school districts while time on science was reduced in 28% of districts and 20% of districts cut recess time. Time on social studies, science, art and music, gym and recess was cut by an average of 145 minutes a week. This amounted to about a one-third reduction in time spent on these subjects.

A RAND study based on surveys in California, Georgia and Pennsylvania found most teachers reporting a narrowing of the curriculum in response to state testing requirements [Hamilton et.al. 2007]. Teachers reported increased time devoted to English and mathematics, most often at the expense of time spent on social studies, arts and music and physical education. A series of case studies of schools in Illinois, Rhode Island and Washington also found that schools had increased time spent on English and mathematics in response to federal and state testing and reporting requirements [Srikantaiah 2009].

Some detailed analytical studies have also found evidence of moving resources from subjects not tested to the subjects that were tested in response to accountability requirements. For example, one study compared trends in mathematics and reading achievement after the introduction of high-stakes testing in Chicago with test score trends in social studies and science, subjects that are not included in the Chicago accountability policy [Jacob 2005]. The gains in mathematics and reading were roughly two to four times larger than gains in science and social studies leading the author to conclude that schools may have shifted resources across subjects.

A recent study of the impact of high-stakes testing in Texas found evidence of strategic resource shifting across subjects within classrooms from those not tested to those that are tested [Reback 2008]. It also found that schools were improving student performance in tested subjects at the expense of performance in other subjects.

Apart from shifting resources between subjects, there is also evidence that schools respond to testing and reporting requirements by more time spent on basic skills teaching and less teaching of more complex skills. Of major concern is that the tests used to measure school performance largely resort to multiple-choice questions and schools increasingly constrict what is taught to skills that are most conducive to testing by multiple-choice.

This approach reduces both the breadth and depth of subject learning. There is a reduction in time spent on exploring more imaginative and creative aspects of the curriculum which are not easily reducible to multiple-choice questions. Shallow learning is also emphasised at the expense of deeper learning and understanding. There is an emphasis on short-term memorisation and 'test tactics' rather than deep learning and understanding

Concerns about the narrowing of subject areas in response to state testing and publication of school results have been expressed in several US studies. For example, a report to the US Education Department noted that there appears to be a widespread tendency for schools to adopt instructional materials and practices that mirror the format of state tests [Hannaway & Hamilton 2008].

In summary, then, there is considerable evidence that the combination of standardised testing and publication of school results leads to a less well-rounded education. The House of Commons inquiry on testing and accountability took considerable evidence on the narrowing of the curriculum and its report concluded:

> ....the way that many teachers have responded to the Government's approach to accountability has meant that test results are pursued at the expense of a rounded education for children. [HC, para 130]

Diane Ravitch, Professor of Education at New York University, who has recently reviewed the experience with testing and accountability programs in the United States concluded:

> When schools are incentivized to measure only basic skills, then everything else loses time and is de-emphasized: the arts, history, geography, civics, science, foreign languages, even physical education. When the test results are used to reward or punish teachers, principals, and schools, then there is even less time for anything that is not tested. When education becomes warped in this way, quality goes down. [Ravitch 2010]

There is already some evidence that publishing NAPLAN school results on My School is narrowing the curriculum in secondary schools in Australia. A survey conducted by the Australian Secondary Principals Association found that 33% of principals said that publication of school results had reduced the breadth of curriculum in their school [ASPA 2010]. Several comments were submitted in the survey about the impact on curriculum. For example:

> …. the effect of the NAPLAN comparisons will be negative. Staff spend more time teaching to the test and less on higher-level thinking skills. The curriculum is narrowed and a school's worth is limited to the average of these 1 hour test scores.
> I have taken periods out of the curriculum in years 7 and 9 to teach to the test. The destructive impact of the simplistic colour coded results of "school effectiveness" is too damaging and has forced this regression to "Educational Fundamentalism" which in itself is shameful. [ASPA 2010: 10]

The survey results indicate that the narrowing of the curriculum is occurring within subject areas. There is increasing emphasis on literacy and numeracy within existing subject offerings and some principals reported that they have reduced the number of subjects offered to students as a result of publication of school test results on the My School website.

3. Distorting teaching practices

Teaching practice also tends to change under the pressure of publishing school results and the pressure to improve league table rankings. Schools and teachers tend to focus more on teaching test-taking skills and practicing for tests at the expense of deeper learning experiences that develop analytical skills and greater understanding.

A widespread phenomenon is “teaching to the test”. This is a term subject to much misunderstanding. It is sometimes observed that so-called “teaching to the test” is not a problem if the test reflects the breadth and depth of the curriculum. However, the problem is that it is difficult to adequately reflect the breadth and depth of the curriculum in a multiple-choice test format of about 30-40 questions to be completed over 45-60 minutes. Schools and teachers emphasise teaching in those areas most conducive to multiple-choice questions and focus on imparting test taking skills and practising for tests.

“Teaching to the test” is better understood as the approach to teaching which emphasises the learning of skills that increase test scores without increasing the underlying skills and knowledge that the test was designed to measure. There are at least two aspects of teaching to

the test. One is drilling of students in the likely questions that they will encounter in the tests through test practice. The other is instructing students in test-taking skills.

In teaching to the test a high proportion of teaching time is given over to test preparation. Teachers coach students on test technique, question spotting, going over sample questions similar to those likely to be set in the test and generally focussing teaching of the substance of a subject in a way best calculated to maximise marks in the test.

There is considerable anecdotal evidence about teaching to the test. For example, a large majority of those who presented evidence to the House of Commons inquiry on testing and accountability were clear that teaching to the test is prevalent and that it is caused by the serious consequences for schools attached to the publication of the results of the tests [HC para 116]. The report concluded that the phenomenon is "widespread" [HC para 130].

Last year, the director of the Oxford University Centre for Educational Assessment and former head of the NSW Board of Studies, Gordon Stanley, warned Australian education authorities about the dangers of teaching to the test. He said:

> We could well end up with a similar situation to the UK, where you get a whole industry created around improving performance on the tests rather than necessarily improving students' learning skills....That has led to a lot of teaching to the test and schools focusing on kids who are close to achieving the targets on the view that they are going to be the easiest to improve. [cited in Patty 2009]

National and state surveys of teachers in the US indicate that teachers in states with high stakes accountability regimes spent a considerable amount on test preparation in class. For example, one survey found that teachers in states with high-stakes tests at the time were much more apt than their counterparts in states with lower-stakes tests to engage in test preparation earlier in the school year; spend more time on such initiatives; target special groups of students for more intense preparation; use materials that closely resemble the test; use commercially or state-developed test-specific preparation material; and use items from previous tests for practice by students [Pedulla et.al. 2003].

Moreover, this emphasis on improving test taking skills is often reinforced by principals and school administrators. For example, a recent study found that about 90% of principals in Georgia and Pennsylvania distributed test preparation materials such as practice tests to teachers, and similar percentages distributed released copies of the state tests [Hamilton et. al. 2007]. Almost all principals said they addressed test preparation in staff meetings and helped teachers identify content that is likely to be on the state test.

This focus on developing test taking skills has reached extreme proportions in Washington DC. In 2008, schools chancellor, Michelle Rhee launched a new program to improve student achievement in Washington DC called the Saturday Scholars program. Under this program, selected students are invited to come to school on Saturdays to work on their reading, math and test preparation skills. In announcing the new program Rhee said its main goal is to "instill lifelong testing competence" [Rhee 2008].

Under this regime, improved education is not about developing lifelong learning, but lifelong testing skills. A broad education and deep learning is not the priority.
Publishing school results and league tables turn classrooms into test preparation factories. In schools all over England and the United States, weeks and months are devoted to test preparation instead of deeper learning. For example, the former head of the UK

Qualifications and Curriculum Agency, Ken Boston, recently commented on the situation in England as follows:

> The amount of time spent on test preparation has increased over the past 10 years: in the second half of the spring term 70 per cent of schools spent more than three hours per week on test preparation. In some extreme cases, months have been spent in the final year of primary schooling on nothing else than test preparation, to the neglect of the other areas of the curriculum and hence to the great detriment of the quality of the children's education. [Boston 2009: 5]

There is considerable evidence that publication of NAPLAN school results on the My School website and publication of school league tables in the media is causing schools to spend more time on test practice in class.

A survey conducted by the Australian Secondary Principals Association found that 65% of schools reported that they had increased the time spent in class on preparation for the NAPLAN tests in 2010 and 70% said they had increased the time spent on practising tests [ASPA 2010].

A report published by the Australian Primary Principals Association said large amounts of valuable instructional time was taken up by coaching and practising tests [APPA 2010]. Some schools spent a significant amount of time on test preparation techniques and allocating test practice as homework to improve their NAPLAN results in 2010. This consumed valuable time that should have been spent on literacy and numeracy and other important areas of the primary curriculum.

There were several reports following the NAPLAN tests in 2010 of schools spending disproportionate amounts of time on test practice to improve school results to the detriment of other areas of the curriculum. The President of the Australian Secondary Principals Association, Sheree Vertigan, said that most students were spending "a lot of time" in class preparing for the NAPLAN tests and that it was "defeating the whole purpose of the tests in the first place" [*The Australian* 10 May 2010]. She said encouraging students to study for NAPLAN tests was "manipulating" the data and was the main catalyst for "narrowing the curriculum".

The Queensland Teachers' Union president Steve Ryan said schools were spending too much time grooming students in an attempt to make them perform better in the NAPLAN tests.

> We've got the ridiculous situation of schools just setting aside a whole range of good curriculum offerings just to concentrate on NAPLAN so the school would be seen to be in some way better than the school next to it. [*ABC News*, 11 May 2010]

The principal of one Perth private school said that it was a "sad week for education":

> The NAPLAN tests are now almost completely useless as a diagnostic tool for individual student performance as different schools choose to prepare so differently, many coaching to the test to avoid potential public humiliation...Stories already abound of a narrowed curriculum filled with NAPLAN lessons from day one, term one from Year 2 onward. [*West Australian* 12 May 2010]

Some schools had prepared for the tests for longer than six months. Numerous teachers told *The Courier-Mail* they spent more than half of their class time on exam preparation and practice questions for NAPLAN [12 May 2010]. The *Canberra Times* reported that some ACT schools had spent up to 90 minutes a day over two weeks practicing for NAPLAN [11

May 2010]. Several parents at a south Canberra school said that students had to sit for two practice tests a day in the lead-up to the NAPLAN tests.

There were also several similar reports about the 2009 NAPLAN tests, the results of which were reported on the My School website in 2010. For example, a report by the Queensland Studies Authority analysing the 2009 test results warned teachers about over-practising for the writing exam. *The Courier-Mail* reported that test markers felt that Queensland students had "over-practised" for the 2009 NAPLAN writing task [11 May 2010]. The *West Australian* reported that up to a quarter of school time was being spent on preparing for the tests and that other subjects such as science had been temporarily dropped to allow more time to practice tests [14 April 2009].

It is reported that many teachers refer to NAPLAN as “napalm” because “it kills everything in the classroom” [*The Courier-Mail,* 18 March 2009].

There is also evidence that school administrators in some states have pressured principals and schools to devote more time to practising for tests. For example, the Victorian Education Department has taken an active role in pressuring principals and teachers to practice for NAPLAN. A memo sent to schools in rural Victoria at the beginning of the 2010 school year told teachers to "explicitly teach" for the national literacy and numeracy tests as part of a drive to lift the state's overall performance [*The Australian* 5 February 2010]. It outlined a 10-week strategy for the department and schools and a "delivery strategy" for teachers in the lead-up to the tests in May. Principals were told to appoint a NAPLAN co-ordinator, set up a sample testing process that may "require resourcing" and "provide additional assistance for students identified as capable of making significant improvement". A "blueprint for classroom approaches" set out strategies to coach students in skills for passing tests, including learning the "test question vocabulary", "skim and scan", and to skip questions that confuse them.

In April last year, the head of the Victorian Department of Education, Peter Dawkins, sent a memo to all principals suggesting more time be spent on preparing students for NAPLAN so as to improve Victoria’s results. The memo stated:

> Students need assistance with preparation to effectively participate in the NAPLAN assessment. They have to understand the genre of testing and the cognitive demands they will be placed under to successfully complete the task. [Dawkins 2009]

The Australian Primary Principals Association president Leonie Trimper said that principals in South Australia were giving consideration to intensive training for students for NAPLAN because they had been placed under extreme pressure from district managers to improve their test results [*The Advertiser*, 30 March 2010].

The president of the Queensland Association of State School Principals, Norm Hartt, said that education officials had told principals to lift their results “at all costs” [*Courier-Mail* 10 June 2010]. Last year, it was also reported that Queensland education officials were putting tremendous pressure on teachers to lift results by practising for tests [*Courier-Mail* 18 March 2009]. Schools were told to put extra time into practising for the tests, even if it was at the expense of time spent on other subjects. The Queensland Teachers Union said that teachers were told that if the results don’t improve, their own employment positions would be reviewed.

The NAPLAN website itself encourages practising for tests. It says that test preparation increases student understanding of what is being asked and provides strategies to focus on test content.

Teaching to the test has even been condemned by Julia Gillard's own deputy chief of staff, Tom Bentley, who was an advisor to the first Education Secretary of the Blair Government:

> The premium placed on test results has encouraged schools and teachers to teach "to the test". This reinforces a system in which students are offered few real incentives to transfer skills across disciplines and contexts or solve real problems within disciplines - to develop their understanding in ways which they could apply in the world beyond the exam hall. [Bentley 2002]

Yet, this is precisely the incentive that his Minister has created for teachers and principals in schools across Australia. It will only do harm to Australian students by restricting and narrowing their learning.

A longer term effect of a focus on test preparation and raising test scores is that it comes to influence ideas about "good teaching". Some studies show that the idea of a "good teacher" changes in schools under the pressure to lift school results. A "good teacher" becomes someone who raises test scores and attains a high pass rate [for example, see Booher-Jennings 2005]. The idea of what it is to be a "good teacher" is itself distorted and corrupted by the pressure to improve test scores.

This development contradicts and endangers the effort to improve teacher quality through recognised standards and professional development. The likelihood is that the ever-present focus on raising test scores will predominate over other notions of "good teaching", and thereby undermine efforts to improve the quality of teaching.

4. Disadvantages low and high achieving students

The pressure to improve school results can also create incentives for schools to ignore students at the lower and upper ends of the student achievement scales. Several English and US studies have found that schools concentrate on improving the results of those students who are on the border of accepted benchmarks and neglect the lowest and highest achievers. This practice is often referred to as "educational triage".

Improving the results of students just below benchmarks is seen as the most efficient way to improve a school's average score or the proportion of students achieving a benchmark. However, while this strategy gives schools a better chance of improving their ranking, it may lead to worse outcomes for low and high performing students.

A couple of anecdotes illustrate how "educational triage" is implemented. According to a *Washington Post* report, the principal of a Washington DC school handed out a list of all the black, Hispanic, special-education and limited-English-speaking students who would take the Maryland School Assessment, the measure of success or failure under the federal No Child Left Behind (NCLB) legislation.

> We were told to cross off the kids who would never pass. We were told to cross off the kids who, if we handed them the test tomorrow, they would pass. And then the kids who were left over, those were the kids we were supposed to focus on. [De Vise 2007]

Those who remained on the list received 45 minutes of intensive test preparation four days a week prior to the tests.

Another anecdote comes from a Texas study where a consultant told teachers:

> Take out your classes' latest benchmark scores …. and divide your students into three groups. Color the 'safe cases,' or kids who will definitely pass, green. Now, here's the most important part: identify the kids who are 'suitable cases for treatment.' Those are the ones who can pass with a little extra help. Color them yellow. Then, color the kids who have no chance of passing this year and the kids that don't count—the 'hopeless cases'—red. You should focus your attention on the yellow kids, the bubble kids. They'll give you the biggest return on your investment. [Booher-Jennings 2006: 757]

Such anecdotal evidence is confirmed by many studies. Booher-Jennings [2005] provides a rich case study of this practice in a Texas school. A recent RAND study of how educators responded to the NCLB accountability requirements in California, Georgia, and Pennsylvania found that some teachers reported focusing more on students near the proficiency cut-off score and expressed concerns about the accountability system's negative effects on the curriculum and instruction provided to high-achieving students [Hamilton et.al. 2007; see also Hamilton & Berends 2006].

A recent study of the impact of the NCLB in the state of Washington found that it provided incentives for schools to focus educational resources on the marginal student rather than those on the tails of the achievement distribution [Krieg 2008]. It found that in schools more likely to be sanctioned under the NCLB requirements, students at either end of the distribution tails make fewer gains than would be expected if their school did not face sanctions. Further, students in the middle of the achievement distribution do better than expected on the high-stakes exam if they attend schools that are likely to be sanctioned.

A study which analysed test scores from 5th grade students in the Chicago Public Schools found that after NCLB was implemented, students performing near the middle of the score distribution showed larger gains than students performing above that level. The gains among the lowest-scoring students were mixed and not as consistent as those of the students scoring near the middle [Neal & Schanzenbach 2010].

Similar results have been found in England. A study using secondary school data found that schools focus on improving the performance of students just below pass benchmarks and that this policy reduced the educational gains and exam performance of very low achieving students [Burgess et.al. 2005].

However, there is some contrasting evidence. Ballou and Springer [2009] found evidence of gains to students at the bottom of the distribution, but find no consistent evidence that schools facing accountability pressure neglect their high achieving students to focus on low achievers. A study using Texas data from the 1990s found that, contrary to the triage hypothesis, when a school has a realistic chance of improving its accountability rating, the lowest performing students make greater than expected gains, even if they have no chance of passing the exam in that subject [Reback 2008]. Nevertheless, it did find some evidence of adverse effects on high achieving students. Another study using North Carolina reading and mathematics test data found little evidence that schools with low results ignore very low achieving students in their efforts to improve their overall school results [Ladd & Lauen 2010]. However, there was some evidence that reading results were reduced for the higher achieving students.

There is also some evidence of in Australia that schools have responded to the higher stakes associated with the NAPLAN test by concentrating on those students most likely to show

improvement if given extra assistance [APPA 2010]. Schools allocated more resources to this select group of students and other students with greater needs did not receive as much attention for the first five months of the year until after completion of the NAPLAN tests.

5. Unfairly stigmatises students and schools

Publication of the results of individual schools and public rankings of schools often unfairly stigmatise and humiliate schools, teachers, students and their families.

The *My School* website uses colour codes to grade schools on their results in four literacy domains and numeracy for a school in each Year level tested. Results substantially below average are to be graded 'red'; those below average are to be graded 'pink'; schools around average will be graded 'white'; those above average will be graded 'pale green' and those substantially above average graded 'green'.

The Federal Education Minister gave an absolute guarantee that *My School* would not "name and shame schools". Yet, this is precisely what it does. Schools with the lowest results are to be flagged 'red', meaning 'red for danger'. It is a signal to parents to keep away. This is public shaming of the worst possible kind.

By 'red flagging' schools, the Government is aiding and abetting an annual ritual hunt for the worst performing schools around Australia that happens in England and the United States. It will unfairly condemn schools in the most difficult circumstances and make their task harder.

Many schools serving the most disadvantaged communities in Australia have been given red flags on My School. They are the schools working in the most difficult of circumstances, yet their reward from the Rudd Government is to be pilloried in public.

Giving schools a 'red flag' is designed to punish. Why else was the colour red chosen for these schools? Punishing schools by publicly labelling them as 'failures' in this way is not the path to school improvement. It is likely to be counter-productive by undermining teaching and learning.

National tests are conducted at two Year levels only in both primary and secondary schools. It means that students in these cohorts are easily identifiable as the "culprits" for a school receiving a 'red flag', especially in small schools, of which there are many in rural areas of Australia. These students will be humiliated and demoralised.

This is not the way to encourage improvement. Students who are humiliated for their learning accomplishments are unlikely to respond positively in their future learning. They could become alienated from schooling and, possibly, from their peers and their community as well.

The task of teachers and schools that are 'red flagged' is made much harder. Far from creating incentives for better performance, 'red flagging' schools is more likely to impair the future learning of some students.

Labelling schools as failures is also likely to set off a spiral of decline. It may cause some parents to leave the school. Reduced enrolments make it harder to retain adequate resources and to keep and recruit good teachers. It could well lead to lower student achievement and lower average school results.

6. Harder for low performing schools to retain teachers

A common perception is that publishing school results, school league tables and other accountability measures make it more difficult for schools with low results to retain and attract quality teachers. While many studies have examined the difficulty high poverty schools have in retaining and attracting quality teachers, only a few have considered the impact of accountability measures in this regard. The weight of evidence in these studies suggests that low performing schools have greater difficulty in retaining teachers.

One study investigated the impact of the introduction of North Carolina's school accountability program [Clotfelter et.al. 2004]. It found that teachers in low-performing schools left teaching at a higher rate in the period following the implementation of the program than in the period before. However, the study did not find evidence that teacher quality was affected by the program. The percentage of quality teachers remained higher in low-performing schools than in middle-to-high-performing schools both before and after the implementation of the program, and no statistically significant differences emerged in the changes that followed implementation.

A review of empirical studies of teacher recruitment and retention made a tentative finding that accountability policies might lead to increased teacher attrition in low-performing schools [Guarino et.al. 2006]. A recent study of the relationship between school accountability and teacher mobility in Florida found strong evidence that teachers are more likely to leave schools that have received a low grading and they are less likely to leave schools that received a high grading [Feng et.al. 2010].

A contrasting result was found in a study of 4th grade teachers in New York State public elementary schools from 1994-1995 through 2001-2002 which examined the response of teachers to the implementation of state-mandated testing. It found that teachers were no more likely to leave schools with low results [Boyd et.al. 2008].

A unique study using data from North Carolina schools examined whether accountability causes reallocation of high-performing principals across schools [Li 2010]. It found that after NCLB was introduced high ability principals at schools more likely to face sanctions for inadequate improvement in student achievement disproportionately move to schools less likely to face sanctions. These changes in the assignment of schools to principals translate into substantive declines in principal effectiveness at schools serving disadvantaged student populations.

7. Discourages co-operation and collaboration

Collaboration is generally seen as an important way to spread innovative approaches and good teaching practice both between and within schools [for example, see Ainscow & West 2006]. Publishing school results and competition for rankings can reduce collaboration between schools and teachers and slow the dissemination of best practice.

First, schools will be reluctant to share successful practices with other schools if it means those schools could leapfrog them in ranking on league tables. For example, a survey by the New Zealand Council for Educational Research found that 38 per cent of schools who faced competition were prepared to share resources or offer mutual support to other schools compared with 58 per cent of schools who did not face competition [Wylie 2006]. A lower ranking for a school could result in a decline in enrolments, less financial resources and greater difficulty in holding and attracting staff.

Second, there is also evidence from overseas that focus on school results undermines co-operation and collaboration between teachers within schools. For example, one detailed Texas case study noted the practice of principals presenting faculty meetings with a "league table" ranking teachers by the percentage of their students passing the latest benchmark test. This engendered a competitive environment within schools and a climate of fear at being labelled a "bad" teacher [Booher-Jennings 2005; 2006]. This put a serious strain on teacher relationships within the school and undermined trust between teachers. Putting teachers into competition with one another led teachers to see their colleagues as threats rather than partners. It served to individualize teacher performance rather than encouraging collective responsibility for all students.

8. Increases social segregation and inequity in education

Publishing school results and publishing league tables tends to increase social segregation and inequity in education. This is driven by both schools and parents. Not only does it increase social division, but it may also harm overall education outcomes, especially for students from low socio-economic status families.

Publishing school results and league tables create greater incentives for some schools to choose their students to maximise their results. Many schools actively choose their enrolments by "cream skimming" students most likely to achieve good results – these students are generally from white, well-off families. Selective and high demand government schools and private schools have control over their enrolments and this allows them greater opportunities to select higher achieving students.

There is abundant evidence from overseas of schools selecting their students to maintain or improve their league table position. For example, a favourite strategy is to use formal and informal enrolment criteria to "cream skim" or "cherry pick" students most likely to achieve good results. In England, the misuse of school admission procedures to maximize school results has been a major ongoing issue since the introduction of league tables. Various methods have been used to select certain groups of students and exclude others, such as giving priority to the children of employees, former students, those with a family connection to the school, and selecting a proportion of children on the basis of aptitude/ability in a subject area(s) or on the basis of general ability [West et.al. 2004]. Many schools also required parents to provide supplementary information unrelated to the school's admissions criteria, such as their occupation, whether the family lived in a hostel or bed and breakfast accommodation and whether parents had refugee status. Many of these practices continue despite government efforts to stamp them out [West et.al. 2009].

A recently published study of public sector religious secondary schools in London has found that selective 'élite' schools appear to 'select out' low income religious families, thereby displacing them to religious schools with a less affluent composition [Allen & West 2009]. It identified a range of different admissions criteria and practices used by the socially selective schools including school-administered banding, aptitude tests, tests of religious or denominational commitment, primary school references and others that may have contributed to the under-representation of lower ability pupils. Moreover, the hierarchy of schools that has developed also tends to dissuade many parents from even applying to the more selective schools because they believe they will not meet the selection criteria.

> ...schools' admissions criteria and practices are important determinants of which pupils are offered places, whether on account of schools attracting applications from certain parents or in terms of the admissions process itself. [Allen & West 2009: 19]

This experience has direct relevance for Australia where private schools have control over their enrolments. Publishing school results and league tables are likely to lead to even greater selection to protect school reputations and rankings. Such selection processes may also extend to the government sector where some schools are in such high demand that they can effectively discriminate amongst applicants for enrolment. Australian governments are unlikely to even attempt to stamp out enrolment discrimination by private and government schools.

On the other hand, many parents see education as a "positional good" in that the value of education depends not on the learning acquired but on the relative status of the school attended. Publishing tables of school results aids this search for status and self-segregation.

In general, it is well-off, white families who make greatest use of choice of school. Extensive evidence from several countries shows that the parents who make use of school performance comparisons and actively choose schools are better educated and have higher levels of income than those who do not [for example, see Cullen et.al. 2005; Hamilton & Guin 2005]. Publishing school results serves these higher income families.

For high income families, finding the "best" schools inevitably means those that have the fewest low-income and minority students. This is demonstrated by a wide range of studies for several cities and regions in the United States and in other countries such as England, Netherlands and New Zealand [see Cobbold 2007]. For example, studies published by the US National Bureau of Economic Research have found that public school choice tends to increase education stratification [Hastings et.al. 2005, 2006, 2009].

A recent study of school choice amongst elementary and middle school students in Durham, North Carolina, compared the characteristics of students who opt out of their assigned school with students who do not [Bifulco et.al. 2009]. It found that well-off families tend to opt out of assigned schools with concentrations of disadvantaged students and to attend schools with higher achieving students. Another study of school choice in Denver, Colorado, found that it was higher income families who were far more likely to participate in inter-district choice and were much more likely to transfer to wealthier school districts. White students were more likely to opt out of racially diverse districts and transfer to districts with greater percentages of white students [Holme & Richards 2009]. Other studies have found that white families tend to choose schools on the basis of their racial composition [Bifulco & Ladd 2007; Saporito 2009]. Another found that parents' decision strategies appeared to be associated with social class status [Bell 2009].

Thus, parents who exercise school choice tend to choose schools with peer compositions that are similar to their own race and social status. By aiding parent choice, publication of school results and league tables can reinforce tendencies toward racial and socio-economic separation in schools.

Increased social segregation in schooling can exacerbate inequity in education in two main ways.

First, it increases disparities between schools in student learning needs and the real resources available to meet those needs. Schools with a high proportion of students from low socio-economic (SES) families have higher levels of learning needs and other problems than high SES schools because low income is associated with lower levels of student achievement. In low SES schools, the scale of challenges is much larger because of the greater concentration of students experiencing them is greater. They generally have to devote far more time and resources to family and health crises, children with few educational materials in their homes, and many children with very weak educational preparation.

Generally, the resources available to low SES schools are not commensurate with the problems they face. Low SES schools are generally funded on the same per capita basis as other schools, with few allowances for the level of need they have to deal with. They have less real resources because they have higher costs and burdens.

Second, increasing concentrations of students from low SES and minority families in some schools tend to lead to lower overall outcomes. Many international studies show that there is a school composition effect on student outcomes associated with high proportions of students from low SES and minority families [see Borman & Dowling 2010; Dronkers & Levels 2007; Oh 2007; Rangvid 2007; Rumberger & Palady 2005, 2006]. A student attending a school where the average SES of the student body is low is likely to have lower outcomes than a student from a similar background attending a school where the average SES of the student body is high. Thus, increasing social segregation between schools is likely to reduce average student achievement.

Furthermore, while the social composition of schools has a significant impact on the achievement of all students, studies show that the impact is greater on low SES and immigrant or minority students [Willms 2006]. There is a "double jeopardy" effect for students from low SES and minority families in that they tend to be disadvantaged because of their circumstances at home, but when they are also segregated into low SES and/or predominantly minority schools they are likely to fare even worse.

Increasing social segregation in schools also has a further effect on the education of the young in society. Schools segregated by class, religion and race make it more difficult for children to develop a real understanding of people of different backgrounds and to break down barriers of social intolerance. Socially segregated schools can feed a lack of social empathy, indeed, social intolerance and an inability of people from different backgrounds to learn to work together and live in the same communities together.

References

Ainscow, M. and West, M. (eds.) 2006. *Improving Urban Schools: Leadership and Collaboration*, Open University Press, Buckingham.

Allen, R. & West, A. 2009. Religious Schools in London: School Admissions, Religious Composition and Selectivity. *Oxford Review of Education*, 35 (4): 471-494.

Alexander, R. (ed.) 2009. *Children, Their World, Their Education*. Routledge, London.

Alexander, R. 2010. The Perils of Policy: Success, Amnesia and Collateral Damage in Systemic Educational Reform. Alexander Miegunyah Lecture, Melbourne Graduate School of Education, University of Melbourne, Melbourne, 10 March. Available at: http://www.education.unimelb.edu.au/news/news%20pdfs/Alexander%20Miegunyah%20lecture%20FINAL.pdf

Australian Primary Principals Association (APPA) 2010. The Reporting and Use of NAPLAN, June. Available at: http://www.appa.asn.au/index.php/appa-business/news-items/993-appa-wants-naplan-reform-media-release

Australian Secondary Principals Association (ASPA) 2010. My School Survey March 2010 Results. Available at: http://www.aspa.asn.au/index.php?option=com_content&view=article&id=369:report-from-the-my-school-impact-on-schools-survey&catid=1:latest-news&Itemid=59

Ballou, D. & Springer, M. G. 2009. Achievement Trade-offs and No Child Left Behind. Paper presented to the *NCLB: Emerging Findings Research Conference* at the Urban Institute, Washington, DC, 12 August. Available at: http://www.caldercenter.org/nclbconference.cfm

Bell, C. 2009. All Choices Created Equal? The Role of Choice Sets in the Selection of Schools. *Peabody Journal of Education,* 84 (2): 191-208.

Bentley, T. 2002. Time to Stop 'Teaching to the Test'. *The Observer*, 10 February.

Bifulco, R., & Ladd, H. F. 2007. School Choice, Racial Segregation, and Test Score Gaps: Evidence from North Carolina's Charter School Program. *Journal of Policy Analysis and Management,* 26 (1): 31-56.

Bifulco, R., Ladd, H. F., & Ross, S. L. 2009. The Effects of Public School Choice on Those Left Behind: Evidence from Durham, North Carolina. *Peabody Journal of Education,* 84 (2): 130-149.

Booher-Jennings, J. 2005. Below the Bubble: "Educational Triage" and the Texas Accountability System. *American Educational Research Journal*, 42 (2): 231–268.

Booher-Jennings, J. 2006. Rationing Education in the Age of Accountability. *Phi Delta Kappan*, 87 (10): 756-761.

Borman, G. & Dowling, N. M. 2010. Schools and Inequality: A Multilevel Analysis of Coleman's Inequality of Educational Opportunity Data. *Teachers College Record*, 112 (5): 1201–1246.

Boston, K. 2009. A Response to the APPA Position Paper on the Publication of Nationally Comparable School Performance Data. August. Available at: http://www.appa.asn.au/images/reports/bostonaddress20090813.pdf

Boyd, D.; Lankford, H.; Loeb, S. & Wykoff, J. 2008. The Impact of Assessment and Accountability on Teacher Recruitment and Retention. *Public Finance Review*, 36 (1): 88-111.

Boyle, B. & Bragg, J. 2006. A Curriculum Without Foundation. *British Educational Research Journal*, 32 (4):569-582.

Burgess, S.; Propper, C.; Slater, H. & Wilson, D. 2005. Who Wins and Who Loses from School Accountability? The Distribution of Educational Gain in English Secondary Schools. Working Paper No. 05/128, Centre for Market and Public Organisation, University of Bristol. Available at: http://www.bris.ac.uk/cmpo/publications/papers/#a2005

Cambridge Primary Review (CPR) 2009. Introducing the Cambridge Primary Review. Available at: http://www.primaryreview.org.uk/

Clotfelter, C.T.; Ladd, H.F.; Vigdor, J.L. & Diaz R.A. 2004. Do School Accountability Programs Make it More Difficult for Low-Performing Schools to Attract and Retain High Quality Teachers? *Journal of Policy Analysis and Management*, 23(2): 251-271.

Cobbold, T. 2007. The Great School Fraud: Howard Government School Education Policy 1996-2000. Available at: http://www.saveourschools.com.

Cullen J. B.; Jacob, B. A. and Levitt, S. D. 2005. The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools. *Journal of Public Economics* 89 (5-6): 729-760.

Dawkins, P. 2009. S172-2009 National Assessment Program – Literacy and Numeracy (NAPLAN) 2009. Memo to Schools and Principals, Department of Education and Early Childhood Development – Victoria.

Dronkers, J. & Levels, M. 2007. Do School Segregation and School Resources Explain Region-of-Origin Differences in the Mathematics Achievement of Immigrant Students? *Educational Research and Evaluation*, 13 (5): 435-462.

Feng, L.; Figlio, D.N. & Sass, T. 2010. School Accountability and Teacher Mobility. Working Paper No. 16070, National Bureau of Economic Research, Cambridge, MA, June. Available at: http://papers.nber.org/papers/w16070

Guarino, C.M.; Santibañez, L. & Daley, G.A. 2006. Teacher Recruitment and Retention: A Review of the Recent Empirical Literature. *Review of Educational Research*, 76 (2): 173-208.

Hamilton, L.S.; Stecher, B.M.; Marsh, J.A.; McCombs, J.S.; Robyn, A.; Russell, J.; Naftel, S. & Barney, H. 2007. Standards-based Accountability under No Child Left Behind: Experiences of Teachers and Administrators in Three States. RAND Corporation, Santa Monica. Available at: http://www.rand.org/pubs/monographs/MG589/

Hamilton, L.S. & Berends, M. 2006. Instructional practices related to standards and assessments. Working Paper WR-374-EDU. RAND Corporation, Santa Monica, April. Available at: http://www.rand.org/pubs/working_papers/WR374/

Hamilton, L.S. & Guin, K. 2005. How Families Choose Schools. In J. R. Betts and T. Loveless (eds.), *Getting Choice Right*. Brookings Institution Press, Washington DC, 40-60.

Hannaway, J. & Hamilton, L.S. 2008. Performance-Based Accountability Policies: Implications for School and Classroom Practices. The Urban Institute, Washington DC. Available at: http://www.urban.org/publications/411779.html

Hastings, J. S.; Kane, T. J. & Staiger, D. O. 2005. Parental Preferences and School Competition: Evidence from a Public School Choice Program. Working Paper No. 11805, National Bureau of Economic Research, Cambridge, MA, November. Available at: http://www.nber.org/papers/w11805

Hastings, J. S.; Kane, T. J. & Staiger, D. O. 2006. Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery. Working Paper 12145, National Bureau of Economic Research, Cambridge, MA, April. Available at: http://www.nber.org/papers/w12145

Hastings, J. S.; Kane, T. J. & Staiger, D. O. 2009. Heterogeneous Preferences and the Efficacy of Public School Choice. Available at: http://aida.econ.yale.edu/~jh529/papers/HKS_Combined_200806.pdf

Holme, J. J., & Richards, M. P. 2009. School Choice and Stratification in a Regional Context: Examining the Role of Inter-District Choice. *Peabody Journal of Education*, 84 (2): 150-171.

House of Commons, Children, Schools and Families Committee 2008a. Testing and Assessment. Vol I: Report, London, 13 May. Available at: http://www.publications.parliament.uk/pa/cm200708/cmselect/cmchilsch/169/169.pdf

House of Commons, Children, Schools and Families Committee 2008b. Testing and Assessment. Vol II: Oral and Written Evidence, London, 13 May. Available at: http://www.publications.parliament.uk/pa/cm200708/cmselect/cmchilsch/169/169ii.pdf

Jacob, B. A. 2005. Accountability, Incentives and Behaviour: The Impact of High-Stakes Testing in the Chicago Public Schools. *Journal of Public Economics* 89 (5-6): 761-796.

Koretz, D.M. 2002. Limitations in the Use of Achievement Tests as Measures of Educators' Productivity. *Journal of Human Resources*, 37 (4): 752-777.

Krieg, J. M. 2008. Are Students Left Behind? The Distributional Effects of the No Child Left Behind Act. *Education Finance and Policy*, 3 (2): 250-281.

Ladd, H. F. & Lauen, D. L. 2010. Status versus Growth: The Distributional Effects of School Accountability Policies. *Journal of Policy Analysis and Management*, 29 (3): 426-450.

Li, D. 2010. Unintended Consequences: No Child Left Behind and the Allocation of School Leaders. Department of Economic, Massachusetts Institute of Technology, Cambridge, MA, June. Available at: http://econ-www.mit.edu/grad/d_li/papers

McMurrer, J. 2008. Instructional Time in Elementary Schools. Centre on Education Policy, Washington DC, 20 February. Available at: http://www.cep-dc.org/

Morton, B.A. & Dalton, B. 2007. Changes in Instructional Hours in Four Subjects by Public School Teachers of Grades 1 Through 4. National Centre for Education Statistics, Institute of Education Sciences, Washington DC, May. Available at: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007305

Neal, D. & Schanzenbach, D. 2010. Left Behind By Design: Proficiency Counts and Test-Based Accountability. *Review of Economics and Statistics*, 92 (2): 263–283.

Oh, G. 2007. Effects of Socioeconomic Status of School Population on Student Math Achievement Growth. Paper presented to the Population Association of America Annual Meeting, New York, 29-31 March. Available at: http://paa2007.princeton.edu/abstractViewer.aspx?submissionId=71932.

Patty, Anna 2009. School Tests Could Fail Students: Educator. *Sydney Morning Herald*, 26 March.

Pedulla, J. J.; Abrams, L.M.; Madaus, G.F.; Russell, M.K.; Ramos, M.A. & Miao, J. 2003. *Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers*. National Board on Education Testing and Public Policy, Chestnut Hill, MA. Available at: http://www.bc.edu/research/nbetpp/reports.html

Rangvid, B. S. 2007. School Composition Effects in Denmark: Quantile Regression Evidence from PISA 2000. *Empirical Economics*, 33: 359-388.

Ravitch, D. 2010. What I Did in June. *Bridging Differences*, 22 June. Available at: http://blogs.edweek.org/edweek/Bridging-Differences/2010/06/what_i_did_in_june.html

Reback, R. 2008. Teaching to the Rating: School Accountability and the Distribution of Student Achievement. *Journal of Public Economics*, 92(5-6): 1394-1415.

Rhee, M. 2008. Fenty, Rhee Launch "Saturday Scholars" Program to Improve Assessment Gains. Media Release, District of Columbia Government, 18 January. Available at: http://edreform.dc.gov/edreform/cwp/view,A,11,Q,460533.asp

Rumberger, R. W. & Palardy, G. J. 2005. Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High Schools. *Teachers College Record*, 107 (9): 1999-2045.

Rumberger, R. W. & Palardy, G. J. 2006. Does Resegregation Matter? The Impact of Segregation on Academic Achievement in Southern High Schools. In J. C. Boger and G.

Orfield (eds.), *School Resegregation*, University of North Carolina Press, Chapel Hill, 128-147.

Saporito, S. 2009. School Choice in Black and White: Private School Enrollment Among Racial Groups, 1990-2000. *Peabody Journal of Education*, 84 (2): 172-190.

Srikantaiah, D. 2009. How State and Federal Accountability Policies Have Influenced Curriculum and Instruction in Three States. Centre on Education Policy, Washington DC, October. Available at: http://www.cep-dc.org/

West, A.; Hind, A. & Pennell, H. 2004. School Admissions and `Selection' in Comprehensive Schools: Policy and Practice. *Oxford Review of Education*, 30 (3): 347-369.

West, A.; Barham, E. & Hind, A. 2009. Secondary School Admissions in England: Policy and Practice. Report prepared for the Research and Information on State Education Trust, London. Available at: http://www.risetrust.org.uk/admissions.html

Willms, J. D. 2006. Learning Divides: Ten Policy Questions About the Performance and Equity of Schools and Schooling Systems. UNESCO Institute of Statistics, Montreal. Available at: http://www.uis.unesco.org/ev.php?ID=6832_201&ID2=DO_TOPIC.

Wilson, D.; Croxson, B. & Atkinson, A. 2006. What Gets Measured Gets Done. *Policy Studies*, 27 (2): 153-171.

Wylie, C. 2006. What is the reality of school competition? Occasional Paper No. 126, National Center for the Study of Privatization in Education, Teachers' College, Colombia University. Available at: http://www.ncspe.org/list-papers.php.