



**National Testing, League Tables and School  
Performance Accountability**

**Prepared by**

**Peter Job  
Federal Research Officer**

**August 2008**

## Contents

Introduction.....	3
Authentic Assessment and Test Validity.....	4
Leaving Children Behind – The US Experience .....	5
Educational Triage – School Performance Accountability in England.....	8
League Tables – A Formula for Failing Our Students .....	10
The “Value Added” League Table Chimera.....	13
Accountability Testing as Ideology.....	16
Accountability Through Teacher Professionalism .....	18
Conclusion .....	21
References.....	25

## Introduction

On 24 July 2007 State and Territory Education Ministers announced a national literacy and numeracy testing program for Years 3, 5, 7 and 9 starting in May 2008. Instigated under the Howard Government, the program enjoys the bipartisan support of all major parties and State and Territory governments. While the program is national in character, states run, mark and report on the tests, and the data stay in their possession. Aggregate State and Territory data are provided for national reporting, with disaggregation for different groups, such as Indigenous, gender and socio-economic status.

Under the present agreement with states and territories, the data are not disaggregated by school and it is not possible to use it for school-by-school reporting or the creation of league tables. It is clear, however, that this situation could change. The Liberal and National Parties have long advocated the publication of school-by-school test data for use for performance monitoring, school-by-school comparisons and league tables. More importantly, it is clear that prominent members of the Federal Parliamentary Labor Party support similar goals. In September 2007 then education spokesperson Stephen Smith said a Labor government would ask the states and territories to publish school league tables based on the tests that compare student achievement in literacy and numeracy. In April 2008 the Deputy Prime Minister and Minister for Education Employment and Training Julia Gillard made it clear she advocated a more comprehensive reporting system to make schools more “accountable for their results” and assure parents received reliable information as to how their schools “perform”.<sup>1</sup>

National standardised testing as a means of school level performance monitoring is flavour of the month, with media commentators, state and federal education ministers and opposition spokespeople talking of accountability, the need to supply information to aid parental choice and the need to monitor supposedly “underperforming” schools. Proponents of national testing claim it will provide a road map to monitor student learning and improvement, a means of making schools and teachers accountable and make them work harder to improve students results. Much of the argument is couched in the vocabulary of social equity, with talk of lifting the performance of underprivileged schools to enhance educational opportunities for Indigenous students and those from lower socio-economic backgrounds.

Whilst the stated aims are laudable, claims that universal test-based performance monitoring programs improve equity and student outcomes are not supported by the evidence. On the contrary, this paper will demonstrate that the weight of international and academic evidence strongly suggests that that a mass standardised testing regime does not deal with the real issues involved in providing excellence and equity in education, and at

---

<sup>1</sup> Tomazin, F. “Gillard to Push for Parents to Get Details on School Performance,” The Age (Melbourne), April 12, 2008. <http://www.theage.com.au/text/articles/2008/04/11/1207856836846.html>

worst is destructive to both quality and equity in education, harming the very students it aims to help.

## **Authentic Assessment and Test Validity**

Effective and authentic assessment lies at the heart of quality education. Authentic assessment is an integral part of curriculum, closely linked to what students learn, providing guidance to students and assisting their learning, evidence of student progress to parents and information to teachers to help them meet the needs of individual students and assess the effectiveness of their own classroom practices. Along with other tools, exams and tests have always been a legitimate and important part of assessment. In particular course-based tests play an important diagnostic role in assessing student aptitudes, achievements and learning deficits. However, effective testing cannot and should not take place in a vacuum. Effective testing must be closely linked to what students are actually studying and undertaken in a manner that is a legitimate and accurate way of measuring what students know and learn.

United States academics Sharon Nichols and David Berliner identify four elements that must be present for a test to demonstrate that it is actually measuring what it is intending to measure, or that it can be said to possess validity.<sup>2</sup> They call this the four Cs: content validity, construct validity, criterion validity and consequential validity. Content validity means that a test should actually measure what it purports to measure; a geology test for example should measure geology knowledge and skills and be based on the actual criteria students were expected to learn in a geology course and curriculum. Construct validity is the extent to which a test measures the attributes and characteristics of the subject it claims to be testing. Criterion validity is the ability of a test to predict and measure a student's achievement now and in the future, that is how they can be expected to perform in a subject at present and as they progress through their education and work. Consequential validity is concerned with the consequences and decisions that are associated with a test score, the use of the data obtained and the extent to which it actually enhances student learning and well-being.

Writing in the US context, Nichols and Berliner are critical of the extent to which externally imposed performance testing is capable of meeting the requirements of any of these criteria, but most particularly content and consequential validity. For a test to possess content validity it must derive from a curriculum and address its needs rather than the other way around. A test imposed upon students that does not derive from the actual curriculum is likely to lead to a distortion of the curriculum itself in order to address the requirements of the test, and this is especially so if high stakes consequences are attached to its outcome. Since any test can only examine a small part of the knowledge and skills in any area, the result may be that of forcing teachers and schools to teach to the test

---

<sup>2</sup> Nichols, S.L. and Berliner, D.C. (2007) *Collateral Damage. How High-Stakes Testing Corrupts America's Schools*. Cambridge, Massachusetts. Harvard Education Press. Chapter 5.

rather than to a broader curriculum, thus narrowing a curriculum and limiting student learning.

For a test to possess consequential validity it is necessary that the information collected be used in a manner that is actually beneficial to students and enhances their learning and as will be discussed, with externally imposed performance testing there are many reasons why this is not always so.

Effective and valid tests must accordingly be diagnostic, written and marked by teachers, assessed against syllabus standards and built around what is actually taught to students. Teacher and school quality cannot be measured by a “snapshot”, especially based on testing divorced from course content. The national testing program claims to be linked to curriculum, but it is not clear how this is so. The same tests are used in all states and territories regardless of the different curricula. While there has been discussion of the development of a national curriculum, if such a system is to be established any national assessment program should be based upon it, not proceed or determine it.

National census testing imposed on the various state curricula will encourage curricula and teaching to be built around tests, and the higher the reporting accountability mechanisms the stronger this trend will be. This will lead to a narrowed curriculum and a lowering of the quality of teaching. There is also strong evidence that high stakes census testing can increase educational inequity, harming the very students it claims to be helping.

The push towards accountability testing in Australia has been strongly influenced by the test-based accountability systems established overseas, most particularly in the United States and Great Britain. It is worth examining these programs and how they have impacted on schools, student achievement and educational equity as a window to the issues surrounding census based accountability testing.

## **Leaving Children Behind – The US Experience**

In 2001 the US Congress and the Bush administration passed the “No Child Left Behind” (NCLB) legislation. Enjoying bipartisan support from both parties, NCLB’s stated aims are laudable, described by President Bush as addressing “the soft racism of low expectations” by narrowing the literacy and numeracy gap between the higher and the lower ends of the educational spectrum, and ending the educational disadvantage of minority students. NCLB requires that all students be tested in maths and reading in grades 3 to 8, and at least once in grades 9 through 12. Federal funding is linked to school performance, with states required to administer their own tests and aim for the ambitious target of 100% proficiency within fifteen years. States are also required to meet mandated state academic standards in core subjects, to publish reports of achievement disaggregated by ethnicity and sub-groups, to publish school league tables based on the testing data and to put into place a series of annual “accountability” based rewards and sanctions on individual schools.

Fundamentally, the NCLB rationale is that of the market, with league tables expected to provide the information to parents to allow them to exercise choice, and sanctions and rewards encouraging achievement by teachers and schools. The latter range in a sliding scale from mandatory funded technical assistance after failure to meet targets after two years, to replacing staff after four and closure to reopen as charter schools after five. In practice commentators have noted that since implementation of NCLB the promised federal funding has effectively disappeared while the punitive measures remain.<sup>3</sup>

Nichols and Berliner document that, as could be expected with such high stakes attached to it, the testing regime has had a profound effect on schools, curriculum, teaching and the treatment of students.<sup>4</sup> It has led to a narrowing of the curriculum, marginalising subjects not tested and narrowing the teaching of even core subjects to areas most specifically related to testing, to the exclusion of subjects such as history and art and aspects of English such as classroom discussions, creative writing and critical thinking. An increasing amount of classroom time is devoted to specifically coaching and cramming for the tests, teaching which has a limited benefit for students' long term educational development, to the detriment of a richer more comprehensive and educationally beneficial curriculum.

Even more seriously, Nichols and Berliner document the fact that NCLB has led to the exclusion from the educational system many of the students it was purportedly designed to help. Students seen as low performing, who were previously seen as challenges schools were obliged to assist, are increasingly viewed as liabilities, and excluding such students from enrolling or encouraging them to leave is an effective way by which schools can meet their proscribed proficiency and improvement targets. Nicholas and Berliner report incidences of students deemed liabilities being bullied into leaving and dropped from rolls in states as diverse as New York and Alabama. Other tactics include suspending students during tests and farming them prematurely into special education streams. The researchers document that high school drop-out rates, particularly amongst minority students, have actually increased in years 11 and 12 since the introduction of NCLB.

According to Nichols and Berliner, high academically achieving students are also adversely affected as schools concentrate upon what they call "bubble kids"; students just behind or on the cusp of achieving the benchmark levels. To maximise a schools' chances of meeting its benchmarks such students are given inordinate priority and resources to improve them to benchmarks levels, to the detriment of both students deemed unlikely to reach them at all, or proficient enough to reach them anyway.

---

<sup>3</sup> Lee, J. (2006) *Tracking Achievement Gaps and Assessing the Impact of NCLB on the Gaps: An In-depth Look into National and State Reading and Math Outcome Trend*. Cambridge, Massachusetts. The Civil Rights Project, Harvard University. p. 5.

[http://www.civilrightsproject.ucla.edu/research/esea/nclb\\_naep\\_lee.pdf](http://www.civilrightsproject.ucla.edu/research/esea/nclb_naep_lee.pdf)

<sup>4</sup> Nichols and Berliner. Op cit.

Other detrimental effects of NCLB documented by Nichols and Berliner include the prevalence of systematic cheating by schools, states and teachers, increased stress levels amongst both teachers and students, erosion of the teacher pool due to problems with retention and recruitment, unfair and unprofessional treatment of teachers in schools deemed “underperforming” and a lack of evidence that NCLB has led to educational improvement.

In 2006 US academic Jaekyung Lee published a report on NCLB under the auspices of the prestigious Harvard University Civil Rights Project.<sup>5</sup> Its conclusions are clear; “This report concludes that neither a significant rise in achievement, nor closure of the racial achievement gap is being achieved.”<sup>6</sup>

Whilst claims to improvement have been made both by states, which have an interest in meeting their mandated targets, and by the Bush administration, which has an interest in depicting its own program in favourable light, the report cites the more independent National Assessment of Educational Progress (NAEP) a long standing national program which has been systematically collecting information since the Nixon administration. The NCLB data demonstrate that progress has not been made since the introduction of NCLB amongst either minority students or the student cohort as a whole. Lee compares this unfavourably with the situation in the 1970s and 1980s, during which substantial progress was made in lowering educational socio-economic achievement gaps under civil rights and anti-poverty programs. The report reiterates the issues of the exclusion of students, the narrowing of curriculum and the demoralisation of teachers and students, characterising NCLB as, “...little more than a theory about how to force change without any grounding in specific educational approaches or targeted resources...”<sup>7</sup>

Citing considerable evidence in its 62 pages that literacy, numeracy or educational equity have not improved under NCLB, the report states in its conclusion; “It is evident that test-driven external accountability, whether it was a state or federal initiative, has not advanced equity on a large scale...”<sup>8</sup>

A more recent, although partisan study of results under NCLB comes from Fairtest, an organisation dedicated to campaigning against what it describes the “deeply flawed” NCLB law and “...to end the misuses and flaws of standardized testing”. Quoting 2007 figures from the National Center for Education Statistics a recent Fairtest paper concludes that improvements in reading and maths scores have either stagnated, slowed, or actually reversed since the introduction of NCLB. While unabashedly partisan, Fairtest is far from a marginal player in the US education debate.<sup>9</sup> As part of a growing national movement

---

<sup>5</sup> Lee. Op cit.

<sup>6</sup> Lee, Op cit. p. 5.

<sup>7</sup> Lee, Op cit p. 7.

<sup>8</sup> Lee, Op cit. p.57.

<sup>9</sup> Fairtest. The National Center for Fair and Open Testing. “*No Child Left Behind*” After Six Years: An Escalating Record of Failure. <http://www.fairtest.org/NCLB-After-Six-Years> Viewed 14 April 2008.

Fairtest has gathered the support of 140 national education and civil rights groups for a joint statement to Congress calling for major changes to NCLB.<sup>10</sup>

## **Educational Triage – School Performance Accountability in England**

Originally introduced under a Conservative government in 1992, National Curriculum tests and their use for the production of school league tables were embraced by Labour when it came to power and extended in 1997 by the establishment of national targets based on school tests. Schools are evaluated by student performance in National Curriculum tests of English, Maths and Science sat at ages seven, eleven, fourteen, and at age sixteen by their five highest General Certificate of Secondary Education (GCSE) results. A more recent addition has been the introduction of a contextual “value added” measure to league tables, taking into account student improvement rather than raw results and factors such as ethnicity and socio-economic background, although Harris notes that these developments have been ignored by the press who continue to use raw test results in their published school league tables.<sup>11</sup>

Many of the same problems observed in the United States under NCLB are identified by critics as operating in England. The National Union of Teachers (NUT) has complained that teachers are forced to teach to the test and that tests distort the curriculum, that undue resources are pumped into students on the margin of the benchmark results to the detriment of others, that league tables confuse parents by measuring factors not related to the quality of teaching and that testing demoralises teachers and students and undermines the quality of education. Furthermore, the NUT and others have pointed out that the result of the National Curriculum Tests school league tables are often considerably at odds with school evaluations produced by the Office for Standards in Education, Children's Services and Skills (OFSTED) system of school inspections which bases its assessment on observed educational practices and qualitative factors rather than test results alone.<sup>12</sup>

Similar issues are also documented by the work of Warwick Mansell in his book, *Education by Numbers – The Tyranny of Testing*. He notes the narrowing of the curriculum caused by accountability testing, the evidence of systemic cheating to enable schools to meet their benchmark targets, the focus on “borderline” students to the detriment of others, and the lack of evidence that the testing and league tables regime has improved educational outcomes, noting an actual decline in English student performance in international measures.<sup>13</sup>

---

<sup>10</sup> Fairtest. Ibid. *Joint Organizational Statement on No Child Left Behind (NCLB) Act*. <http://www.fairtest.org/joint-organizational-statement-no-child-left-behind> (Viewed 14 April 2008)

<sup>11</sup> Harris, B. (2007) *Why ranking Schools Would do More Harm than Good*. Paper for the AEU Federal Executive, Melbourne. p.5.

<sup>12</sup> National Union of Teachers. *The Case Against National Curriculum Tests*. [http://www.teachers.org.uk/resources/pdf/case\\_against.pdf](http://www.teachers.org.uk/resources/pdf/case_against.pdf) (Viewed 22 April 2008). Information also cited in Harris, Ibid. p. 6.

<sup>13</sup> Mansell, W. (2007) *Education by Numbers. The Tyranny of Testing*. London. Politicos Publishing.

The exclusion of lower performing students from schools has also been documented. In a study of school selection practices amongst London comprehensive schools West and Hind found that those with autonomy over their own admissions were more likely to exclude children with special social, medical and educational needs than schools whose admissions were controlled by a local authority, and that they did so on the basis of the effect this would have on their standing in the league tables.<sup>14</sup> As could be expected, schools with such policies had lower numbers of students from lower socio-economic and immigrant backgrounds, scored higher in public testing regimes and league tables and found it easier to reach their benchmark targets. West and Hind accordingly attribute the steep hierarchy in school positions in league tables in London to these school admission practices.<sup>15</sup> Other writers have documented evidence of lobbying of politicians by principals of non-selective schools to select their enrolments in order to improve their rankings in league tables.<sup>16</sup>

University of London academics David Gillorn and Deborah Youdell conducted an extensive study on the effects of National Curriculum Tests and league tables on two English government schools, concluding that they deepened racial, ethnic and class inequalities and led to irrational and unfair allocation of educational resources.<sup>17</sup> Furthermore, they claim that students are allocated educational resources in a similar manner in which US schools invest disproportionately in “bubble” students; those on the cusp of achieving the required test results to meet the school benchmarks are helped to the detriment of those below or above. Gillorn and Youdell label this process “educational triage”;

“Here we see the first element of educational triage: resource allocations are managed so that additional attention is limited once it is clear that someone will be all right. It may be painful, but the situation is worse for those designated as beyond hope.”<sup>18</sup>

Gillorn and Youdell note that educational triage excludes lower performing students from the assistance they need, and also has a detrimental effects on academically achieving students whom “...place demands on the school that are at odds with, that is in excess of, the demands of Government and educational bureaucracy.”<sup>19</sup> In one of the schools they studied, for example, higher achieving students were discouraged from taking more than the five minimum GCSE subjects they needed to meet the school performance

---

<sup>14</sup> West, H. and Hind. H. (2006) *Selectivity, Admissions and Intakes to “Comprehensive” Schools in London, England*. Educational Studies, Vol. 32, pp. 145 – 155.

<sup>15</sup> Ibid. p. 153.

<sup>16</sup> Rowe, K.J. (2000) *Assessment, League Tables and School Effectiveness: Consider the Issues and ‘Let’s Get Real!’* Journal of Educational Inquiry, Vol.1 No.1. p. 76.

<sup>17</sup> Gillborn, D. and Youdell, D. (2000) *Rationing Education*. Buckingham. Open University Press.

<sup>18</sup> Gillborn, D. and Youdell, D. (1998) *Raising Standards and Deepening Inequality: Selection, League Tables, and Reform in Multiethnic Secondary Schools*. Paper presented at the symposium, “Racism and Reform in the United Kingdom: The Market, Selection, and Inequality,” held at the Annual Meeting of the American Education Research Association. (San Diego, CA, April 1998). p. 12.

<sup>19</sup> Ibid. p.12.

benchmarks to allow school resources to concentrate on those students on the cusp of the required results.<sup>20</sup>

Gillborn and Youdell also discuss one of the rationales advanced by the supporters of league tables, that they aid parent choice. They note that while parents have a right to express a preference for a certain school they have no right of access. Schools which listed higher on the league tables are accordingly in a position to choose their students, and given they are likely to do so in the interests of maintaining their positions in league tables this has a further detrimental effect upon students from lower socio-economic and minority backgrounds.<sup>21</sup>

The researchers also note the increasingly narrow use of the term “standards” by policy makers under the National Curriculum Testing regime. Increasingly it denotes measurable outcomes in externally examined tests rather than more comprehensive concepts of educational achievement, further narrowing the concept of what an effective curriculum should be in the minds of policy makers.

It should be noted that the performance table regime no longer extends to the whole of Britain. In the face of the issues discussed Northern Ireland and Wales abolished league tables in 2001 and Scotland in 2003. In 1998, Welsh children were lagging behind English children on almost all testing and examination indicators. But by 2002 Welsh children were doing as well or better than the English without the same amount of targets, tests and league tables seen as essential to success in England. This success is being produced despite higher levels of social deprivation than England but spending, pro rata, a similar amount of money on education.

## **League Tables – A Formula for Failing Our Students**

In response to the Howard government proposals for national testing leading to school level performance reporting and league tables the Australian Capital Territory Minister for Education and Training commissioned the ACT Government School Education Council to produce a report on the Federal government proposals, and in June 2004 the council delivered its report, *School Performance Information*.<sup>22</sup>

The council’s conclusions were clear. They reported that reporting individual school results as a measure of school quality, whether raw or value added, is unreliable, misleading, and damaging to schools and the quality of education they deliver. They concluded that the use of exam and test scores to judge schools publicly is not useful either as a device to assist parents in choosing a school or as a form of public

---

<sup>20</sup> Ibid. pp. 11 – 12.

<sup>21</sup> Ibid. p.3.

<sup>22</sup> *School Performance Information. An Issues Paper for the ACT Minister for Education and Training.* (2004). Government School Education Council, ACT.

[http://www.gsec.act.edu.au/\\_data/assets/pdf\\_file/0009/38808/School\\_Performance.pdf](http://www.gsec.act.edu.au/_data/assets/pdf_file/0009/38808/School_Performance.pdf)

accountability, and as such are unable to deliver the objectives of the Australian Government.

Furthermore they stated;

“The Government Schools Education Council (A.C.T) considers that the Australian Government proposal is likely to undermine a key principle and policy goal of public education in the ACT, namely, improvement of equity in school outcomes.”<sup>23</sup>

The report summarises the arguments against league tables in a series of dot points;

“League tables are likely to:

- exacerbate the problems of misleading and inaccurate information about school performance associated with reporting school averages;
- lead to a public debasement of schools with very poor results and a low ranking, and to public labelling of their students and families as ‘failures’;
- lead to greater education inequities and social segregation of schools as high ranking schools select ‘good’ students and reject ‘poor’ students, while the best teachers move to the high ranking schools; and
- undermine effective school improvement.

“League tables may provide misleading rankings of schools because:

- rankings of raw results tend to reflect the socio-economic background of the families of students;
- different measures of school results lead to different rankings of schools; and
- rankings can create non-existent differences between schools and create failure where none exists.”<sup>24</sup>

The very nature of league tables, and the public reporting of data creates a situation of blame and increases inequities. A market based concept of competition inevitably creates winners and losers, designating schools as successes and failures. This is true whether or not the data is true or misleading, and league tables make it inevitable not only that some schools will be stigmatised, but often stigmatised regardless of the actual quality of education delivered or the commitment of their teachers. As the ACT report points out, the labelling of some schools as “failing” creates a climate of recrimination and retribution which may undermine teaching and learning.<sup>25</sup> The report quotes several sources to demonstrate that under such a process students become very much aware of the status of their school, and that this has negative consequences for their own self images and commitment to education;

“Students who are humiliated for their learning accomplishments are unlikely to respond positively in their future learning.”<sup>26</sup>

---

<sup>23</sup> Ibid. p. 25.

<sup>24</sup> Ibid. p. 9.

<sup>25</sup> Ibid. p. 12.

<sup>26</sup> Ibid. p. 13.

The stigma attached to schools labelled as “failing” and “underperforming” can damn a school in the eyes of the public and parents, causing parents to withdraw students from a school or decline to enrol them regardless of the actual quality of education and school programs. Supporters of the public publication of school “performance” data see this as an appropriate and positive encouragement of the exercise of parent choice, yet not only is such choice often based on misleading information but the actual choice exercised is constrained by the selection policies of schools. As previously discussed, schools feel obliged to select students in the interests of their own “performance” rankings and school image, excluding those they deem a liability and further limiting such students’ choices. Schools with higher rankings effectively choose students rather than parents choosing schools, leading to the further narrowing of the educational options for students from underprivileged backgrounds and creating greater societal segregation and educational inequities. As the ACT report notes, in England there has indeed been an increasing segregation of schools by both class and race since the introduction of league tables.<sup>27</sup>

When schools are labelled as underperforming, teachers in such schools run increased risks of being stigmatised as failures regardless of their commitment or capabilities. Due to the increased stigma and stress, as well as the potential damage to their careers in a blame-based accountability system, teachers are likely to become increasingly unwilling to work in disadvantaged schools.

By its nature a testing based performance regime discourages school collaboration and leads to an increasing focus on the marketing of a school image rather than educational improvement. To fit the tests schools reduce both the scope of the curriculum and the nature of their teaching. The narrowing of the curriculum caused by an excessive emphasis on performance measures is demonstrated by a quote from a teacher in England, originally from the Daily Telegraph, reported by Warwick Mansell;

“Securing good coursework results is simple. All my pupils ‘chose’ the same questions. I teach them lessons in which I more or less dictate the answers. I give them worksheets with detailed guidance on how to answer each question. I give them words with which to start and link their sentences. I mark their work, correct the grammar and then re-mark it until they get it right. I hate doing it. But it works.”<sup>28</sup>

Both lower and higher academically capable students suffer as their educations and needs are subjugated by the requirements for a school to meet performance benchmarks based on narrow accountability measures.

On 12 April 2008 Education Minister Julia Gillard told the media that she would push the states for the publication of information on how schools perform based on results from the literacy and numeracy tests to allow parents to know “...what’s happening in their

---

<sup>27</sup> Ibid. p. 16.

<sup>28</sup> Mansell. Op cit. p.77.

local school.”<sup>29</sup> Although she described league tables as “...not an appropriate term”, experience indicates that once such data is released newspapers will quickly tabulate and publish them as league tables.

The extent to which the Federal Government will succeed in this goal is unclear. State and Territory positions on the publication of disaggregated school testing data varies across the spectrum, from NSW which has specific regulations against release, to West Australia which effectively already allows it. It is clear, however, that the existence of national census testing regime will lead to an inevitable impetus towards the release of data and the publication of league tables.

The ACT report argued that the release of information that could seriously harm individuals or mislead the public is not always in the public interest, and that, “...individual school results and comparisons of results in league tables are certainly in this category.”<sup>30</sup> There is also a strong argument that student performance data ethically belongs to students, teachers and parents and should not be released publicly in ways that could damage and stigmatise students and schools.

The creation of school league tables would lead to a diminishing of the quality of education in Australia, a narrowing of the curriculum to the detriment of all students, and an abandonment of goals of education to increase social equity. It would increase privilege, inequity and disadvantage and harm the very students the supporters of such as regime most claim to wish to help.

## **The “Value Added” League Table Chimera**

As many commentators have noted, raw school averages usually indicate more about a school’s cohort than its educational practices. In response to the obvious irrationalities and draw backs of league tables based on data from different schools in different situations with different student populations, some have advocated the use of “value added” performance measures which take into account students prior achievements and socio-economic elements of a school’s cohort. Such a response has found favour in Australia. The pre-summit discussion document relating to the productivity and education section of the 20/20 Summit states;

“Teacher quality will arguably become better understood when analysed in the emerging context of ‘value adding’ practices in education. This is the use of increasingly sophisticated measures and technologies to monitor the progress of students within a ‘similar ability,’ or with the rest of the state.”<sup>31</sup>

---

<sup>29</sup> Tomazin, F. “Gillard to Push for Parents to Get Details on School Performance,” *The Age* (Melbourne), April 12, 2008. <http://www.theage.com.au/text/articles/2008/04/11/1207856836846.html>

<sup>30</sup> *School Performance Information*. Op cit. p. 24.

<sup>31</sup> Australia 2020 Summit. (2008) *Education Skills and the Productivity Agenda*. Background paper p. 10. <http://www.australia2020.gov.au/topics/docs/productivity.ppt#258.1> Viewed 8 April 2008.

On April 13 2008 Julia Gillard commented on the need to give parents more information on school “performance” and the “need to understand in a much more sophisticated way what’s going on in schools”, including measures of how much schools “add value” to student performance.<sup>32</sup>

The concept of value added comes from business economic models, as measuring the value added by a business or process to a product or service. Value added methodologies are presently in use in several states in Australia including Tasmania and New South Wales for purposes such as tracking the progress made by students, determining for internal department use the needs of individual schools, and for longitudinal studies of student achievement. The effectiveness of value added methodologies overall remain a matter of debate within academic circles between those who regard it as an effective mechanism and others who are more sceptical.<sup>33</sup> Whatever its validity for academic or internal departmental purposes, however, there is strong evidence that as means of making public comparisons between the performance of different schools the value added methodology is more likely to create further confusion and perpetuate stereotypes than address the obvious drawbacks of assessing comparative school performance on raw test score data alone.

Value added performance measures were introduced in England in 2002. In the face of mounting challenges to the validity of league tables and raw test data, Wales, Scotland and Northern Ireland chose to abandon publishing data altogether. England instead chose to try to preserve the census testing and league tables approach by judging schools not on their raw data alone, but by how much they supposedly “added value” to their cohorts and boosted their performances, judged by their primary school scores. The validity of such measures, however, has been questioned by researchers on a number of grounds.

Writing in 2005, Stephen Gorard of the University of York examines the 2004 scores based on the value added concept in York, Leeds, East Riding or Yorkshire and North Yorkshire. He found that schools that produced a high value for raw “performance” scores also produced a very similar high score for the “value added” score as well, and vice versa. There are no low or mid attaining schools with high value added scores, and no high attaining ones with low ones. In fact he goes as far as saying that “...we could predict the value added figure for any school extremely well just from their absolute level of final attainment.”<sup>34</sup> Gorard concludes that so-called value added scores are measuring the same things raw scores measure, namely the underlying socio-economic factors that make the “performance” scores for different schools with different cohorts different in the first place;

---

<sup>32</sup> Tomazin, F. “Gillard to Push for Parents to Get Details on School Performance,” *The Age* (Melbourne), April 12, 2008. <http://www.theage.com.au/text/articles/2008/04/11/1207856836846.html>

<sup>33</sup> Academics Doehke, Locke and Petrosky, for example, describe value added methodology as a “crude” mechanism of limited value. See Doehke, D. Locke, T. and Petrosky, A. (2004) *Explaining Ourselves (To Ourselves): English Teachers, Professional Identity and Change*. Literacy Learning, Issue 139, pp.103-112.

<sup>34</sup> Gorard, S. (2006) *Value-Added is of Little Value*. *Journal of Education Policy*, Vol. 21, No. 2, p. 239.

“Value-added scores are no more independent of raw-score levels of attainment than outcomes are independent of intakes.”<sup>35</sup>

Writers in the US have also pointed out that value added school performance scores can be affected by elements such as student mobility, changes in the demographic profile of schools and a variety of similar factors. Measurement factors such as statistical uncertainty and sampling errors can also cause fluctuations from year to year. Kane and Staiger from the National Bureau of Economic Research in Cambridge, Massachusetts conclude;

“Although gain scores are often touted as better indicators of a school’s ‘value added’, they are much more likely to be affected by idiosyncratic fluctuations in scores from year to year...attempting to estimate a school’s value-added is analogous to looking for a smaller needle in a bigger haystack.”<sup>36</sup>

Linn and Haug of the University of California reach similar conclusions;

“This volatility results in some schools being recognised as outstanding and other schools identified as in need of improvement simply as the result of random fluctuations.”<sup>37</sup>

Since 2006 England has added a Contextual Value Added (CVA) measure that takes into account a cocktail of socio-economic factors including poverty, ethnicity, English as an additional language and special educational needs. However, in January 2008 a study by Bristol University received considerable media attention by casting doubt on the effectiveness of such measures. The researchers point out that nearly half of all secondary schools are judged around average in the value they add. Therefore, the scores of many schools are very similar, leading the differences in their positions on league to be determined by very small differences in scores. These small differences, far too small for reasonable comparative differences in educational quality to be deduced, nevertheless leave them hundreds of places apart in the tables.

“...our analysis highlights further major problems with the current use of CVA to rank and compare individual schools. Given that almost half of the schools cannot be statistically distinguished from the national average, any ranking exercise based on these numbers will be largely spurious.”<sup>38</sup>

---

<sup>35</sup> Ibid. p. 240.

<sup>36</sup> Kane, T.J. and Staiger, D.O. (2001) *Improving School Accountability Measures. Working Paper 8156*, National Bureau of Economic Research, Cambridge. Cited in, *School Performance Information. An Issues Paper for the ACT Minister for Education and Training*. Government School Education Council, ACT p.7. [http://www.gsec.act.edu.au/\\_data/assets/pdf\\_file/0009/38808/School\\_Performance.pdf](http://www.gsec.act.edu.au/_data/assets/pdf_file/0009/38808/School_Performance.pdf)

<sup>37</sup> Linn, R.L. and Haug, C. (2002) *Stability of School Building Accountability Scores and Gains, CSE Technical Report 561*. Center for the Study of Evaluation, National Center for Research and Evaluation, Standards and Student Testing, University of California, Los Angeles. Cited in *School Performance Information*. Ibid. p. 7.

<sup>38</sup> Wilson, D. and Piebalga, A. (2008) *Accurate Performance Measure but Meaningless Ranking Exercise? An Analysis of the English School League Tables*. The Centre for Market and Public Organisation. Bristol Institute of Public Affairs, University of Bristol. p. 11. <http://www.bris.ac.uk/Depts/CMPO/workingpapers/wp176.pdf>

In comment reported by the Guardian newspaper the researchers were franker still, claiming that such contextually value added tables are “meaningless”, with some of the data “at best misleading, at worst dishonest.”<sup>39</sup>

Whilst the concept of being able to make a precise measure of the value a certain school in a certain time adds to a student’s education may be appealing, the complexities involved make this impossible in practice. Education is a qualitative experience over time that is different for every student and for different cohorts, encompassing a wide range of factors that simply cannot be measured in this way. Far from being the sophisticated measures their proponents claims them to be, external test-based “value added” measures used for overly simplistic school or teacher comparisons are likely to confuse a situation even more by adding an apparent veneer of evidence based substantiation to what are fundamentally flawed determinants of school or teacher quality. As Gorard puts it;

“In fact, the value-added calculations are rather worse than pointless because their apparent precision and technical sophistication may have misled analysts, observers and commentators into believing that they had succeeded, or that a greater range of variables or a more complex analytical model would somehow solve the outstanding problems.”<sup>40</sup>

The implications for value added league tables for public schools serving underprivileged cohorts are clear; they would continue to get labelled as “failing” schools under the value added model as much as they would using raw scores alone, except that their failure would supposedly be determined by a more “sophisticated” measure, and be assumed to be more “scientific” and “evidence based”. Even if this were not so, it is likely that public and high stakes school performance measurements would continue to be damaging under a value added model. Other negative consequences would continue; the exclusion of students and cohorts perceived as harming a school’s standing; the issues around teaching to the test and the narrowing of the curriculum; the failure to cater to the diverse needs of academically capable students; the failure to value and invest in diverse teaching strategies that encourage critical thinking, challenge students and leave with them with the flexible range of skills twenty-first century society and workplaces require.

## **Accountability Testing as Ideology**

Nichols and Berliner note that one of the major impetuses towards testing and league tables has been the increasing prominence given to business accountability models throughout society.<sup>41</sup> Such models seek ways in which to monitor and increase productivity without spending more money. The application of such a “productivity” model to the study of educational outcomes accordingly demands supposedly numerically measurable performance outcomes to be measured against financial expenditure. Testing is accorded an important place in such an approach as it produces the most easily

---

<sup>39</sup> Shepherd, J. “ ‘Fairer’ School League Tables Misleading – Study.” The Guardian, (Manchester) January 7, 2008. <http://www.bris.ac.uk/Depts/CMPO/workingpapers/wp176.pdf> Viewed April 22.

<sup>40</sup> Gorard, Op cit. p. 241.

<sup>41</sup> Nichols and Berliner. Op cit. p.18.

measurable data. Such an approach, however, simply does not account for the range of outcomes and diverse factors that contribute to a modern education. It does not measure important qualitative factors that are important to an individual's participation in modern society, her or his ability to acquire the skills necessary to survive in the modern workplace, nor does it sufficiently take into account the complexity of "inputs" involved in diverse student cohorts over time.

Those in support of testing accountability models often talk of an evidenced based approach. By this they do not usually mean an approach based on the professional knowledge of teachers or educators, nor the consensus of academic research around the world as to what constitutes effective educational practice. Instead they usually mean often simplistic attempts to measure school and teacher performance through metrics. This leads to the need to test, to measure and compare, to create league tables, to invent value added mechanisms when others are discredited, to produce complex numerical calculations of school and teacher performance. Metrics can deliver a false sense of sophistication, a claim to an "evidence based" scientific methodology, whether or not the validity of such an approach is actually supported by evidence. It is ironic in this context that so much evidence on the damaging effects of census testing and league tables is blithely ignored by those supposedly demanding evidence based approaches.

Market based accountability models by their nature emphasise competition and judge successes and failures. Such an approach is accompanied by a strong sense of blame accorded to those perceived to be the latter, whether schools, teachers or students, along with the notion, largely developed by non-teaching "experts", that educational improvement can be obtained not through resourcing and support, but by surveillance and performance mechanisms based on supposedly measurable data. The public education system, which teaches the majority of students from underprivileged backgrounds, is particularly seen in this light, and teachers within it are often viewed not as professionals whose opinions should be valued but as underperforming obstacles to change. The political and media vilification of what are frequently described as "underperforming" schools and teachers can be understood from this perspective.

A business accountability model is also naturally averse to the spending of public money, and implementing a testing regime is cheaper than putting in the resources needed to address educational disadvantage. Research by the Ministerial Council of Education, Employment and Youth Affairs (MCEETYA), shows that public schools nationwide are under-funded to the tune of \$2.9 billion per year. Yet even in this context there are those who claim that educational improvement in underprivileged public schools can be achieved not by resourcing and support but by surveillance, testing, mandatory benchmarks, naming, shaming and punishment.

Such an approach has a particular appeal to policy makers with an eye on the budget bottom line. In April 2008 the Victorian government invited Sir Michael Barber, one of the chief architects of the English school performance monitoring regime, to address by teleconference a summit of senior government and education figures in support of the state's own educational agenda. Sir Michael stated that "...just spending money – more

and more of it – doesn't work..." and that nor does "...simply reducing student-teacher ratios, or simply increasing spending per student...". A self proclaimed supporter of league tables he instead emphasised the need to improve "the quality of teachers", claiming that socio-economic educational differences could be addressed by "a highly effective educational system", apparently one modelled upon the kind of performance monitoring regime in place in England.<sup>42</sup>

As a teacher familiar with the dilapidated classrooms and underfunded programs prevalent in Victorian public schools, this writer's reaction to Sir Michael's remarks was an overwhelming sense that it was not the message needed by the lowest education spending state of one of the lowest education spending countries in the OECD. There was also some bemusement that one of the foremost architects of a regime with such demonstrably poor results should be consulted as an authority. Nevertheless such opinions were taken to heart by the Victorian Education Minister Bronwyn Pike who eight days later stated on television;

"All of the international research in fact shows that it's not necessarily extra money that is the really critical factor in turning around underperforming education systems. It's really about the quality of the leadership and the shared professional development agenda that makes the difference."<sup>43</sup>

Given her adherence to the views of Sir Michael and previously stated commitment to dealing with "underperforming teachers", the minister's "professional development agenda" can most accurately be read as performance accountability mechanisms.

Blaming schools and teachers for real and perceived educational deficiencies is an effective mechanism for deflecting scrutiny away from important issues of educational inequity and the under-resourcing of public education. The present emphasis on school and teacher accountability shifts focus away from the responsibility of governments to assure that all students have access to a quality education. Accountability, particularly for the education of underprivileged children, is increasingly being placed at the feet of public schools and their teachers alone and conflated with acceptance of assessment and reporting initiatives based upon external standardised testing. Yet true accountability has always been about far more than this and alternatives are available.

## **Accountability Through Teacher Professionalism**

The insights of outside observers of the Australian education system can often put educational debates taking place within the country into some perspective. In an address in 2007 to the European Reading Conference in Berlin, Andreas Schleicher of the OECD noted that Australia's Program for International Student Assessment (PISA) literacy

---

<sup>42</sup> Tomazin, F. "Pouring Money Into Schools 'Doesn't Work'," *The Age* (Melbourne), April 11, 2008. <http://www.theage.com.au/news/national/pouring-money-into-schools-doesnt-work/2008/04/10/1207420591367.html> (viewed 11 April)

<sup>43</sup> Stateline Victoria. ABC Television. 18 April, 2008. Transcript - <http://www.abc.net.au/stateline/vic/content/2006/s2221256.htm> (viewed 24 April 2008)

results placed it amongst the top tier of countries. He went on to attribute this high level of literacy achievement by Australian students to a high level of Australian teacher professionalism.<sup>44</sup>

School and teacher accountability for the quality of education they provide is essential. It is a concept supported by the great majority of schools and teachers, and by the Australian Education Union. Parents, students and the public have a right to know that schools and teachers are competent, professional, and engaged in high quality teaching practices. There is, however, a very great deal of difference between an accountability system based on supporting teachers and recognising their professionalism and one based on mistrust, blame, unreasonable sanctions and the use of inaccurate data or factors beyond teachers' control as assessment tools. Nor should accountability end at the classroom door. For schools and teachers to be properly and appropriately accountable and deliver the best education possible they require the support of wider education systems, governments and the community.

In 2007 then AEU research officer Roy Martin wrote a paper for the AEU examining the concept of teacher accountability. It is worth quoting at some length;

“The quality of teaching is undoubtedly extremely important, and this is borne out in the research. However, the implications from this take two very different paths depending whether this leads to greater or lesser trust. The high trust path puts an emphasis on measures which enhance judgment, respect expertise, encourage collegiality, ensure participation in decision making, offer system support and are tied to high status and rewards for all. It acknowledges that teaching and learning is problematic and that there is a need to search for answers on these problems. The route to higher quality is seen to be through research, professional learning, dialogue and learning from each other.

“The low trust path imposes accountability mechanisms designed to monitor and regulate. It tells teachers what to teach and how to teach it, and then introduces testing to measure how well this is happening. If students have problems, it is perceived to be because teachers are not doing their job properly. Teachers and schools with achieving students get rewarded, whilst those struggling with the most problems get blamed for failure. The route to higher quality on this path is perceived to be mandated curriculum developed by non-teaching “experts”, national testing, public reporting of results, performance pay, incentives for “success” and punitive measures against “under performing” teachers and schools.

“It is the medical equivalent of arguing the reason people die of cancer is because doctors are not up to the task and switching the emphasis from research on the

---

<sup>44</sup> Schleicher, A. (2007) *Literacy Skills in the Information Age*. Paper presented to the 15<sup>th</sup> European Conference on Reading, Humboldt University, Berlin. Quoted in Luke, A. and Woods, A. (2007) *Learning Lessons” What No Child Left Behind can teach us about literacy, testing and accountability*. Australian Curriculum Studies Association.

[http://www.acsainc.com.au/content/lessons\\_from\\_no\\_child\\_left\\_behind.doc](http://www.acsainc.com.au/content/lessons_from_no_child_left_behind.doc)

causes and cures of cancer to developing carrots and sticks to make doctors work better.”<sup>45</sup>

Whilst the quality of teaching and the professional standards of teachers are clearly very important factors in the educational outcomes of students, teachers cannot achieve their best in isolation, and adequate support is necessary to allow them to achieve their best. Support for public education, moral as well as financial, is essential to addressing educational inequity and lowering the differentials in educational outcomes. While all schools should be subject to reasonable and credible accountability mechanisms, they should be supported with proactive and collaborative assistance rather than ill informed and paternalistic surveillance and intervention strategies. Within this context, valuing teacher professional judgements is critical for student growth and achievement and essential to a fair and effective accountability system. An accountability model based on teacher trust would ensure that teachers have the resources and opportunities to apply their professional judgements to their teaching.

The primary purpose of an accountability system should be to support student learning. Teachers accordingly have a number of responsibilities; to possess and regularly update their knowledge of their subjects; to use it to make professional judgements to enhance student learning; to work collaboratively with other professionals, with the school and the wider education system to produce effective programs and make the best decisions for their students; to explain and justify their decisions about student learning to students, parents and the public. Within this context, accountability for teacher quality and professionalism must be based upon independent peer based qualitative appraisal that not only assesses teacher effectiveness but provides support for improvement.

Teacher accountability accordingly must be accompanied by systems accountability. Governments have a responsibility to assure that all schools and teachers are provided with the resources to do their jobs properly, and that resources for additional intervention programs are provided where situations of special needs arise. This implies an obligation to find ways to improve the capacity and performance of those responsible, not just measure the achievement of outcomes. This includes improving pre-service education programs, providing strong support to beginning teachers, investing in ongoing professional development to assist teachers throughout their careers, and providing the resources to ensure that there is adequate time for teachers to take advantage of opportunities for professional development. Proactive safeguards should also be put in place to encourage good practices, prevent abuses, ensure some course of redress for problems that arise and provide some assurance of equitable and fair treatment.

Some claim that national literacy and numeracy testing programs provide the information to identify students and cohorts with special needs and allow such resources to be provided. Most of the information, however, is already available to education departments and schools. Where further diagnostic information is required it can be obtained through the use of sample testing. Sample testing does not lead to “teaching to the test” nor the distortions of the curriculum caused by full cohort testing, and is

---

<sup>45</sup> Martin, R. (2007) *Low Teacher Trust – Let’s Not Go There*. AEU document.

accordingly a more accurate way in which to assess standards. As the 2007 AEU Curriculum Policy contends, sample testing can provide the system-wide information required to support planning and resource allocation and enable governments and education systems to fulfil their responsibility to provide funding for programs in areas identified as in need.<sup>46</sup> This should take place in the context of discussion with teachers, students, education unions and parent groups about proposed use of such data, and with strict protocols for privacy to assure that public reporting on individual students and schools does not take place and the information cannot be used for the construction of league tables.

Accountability has to be with governments and education departments to use information responsibly to enhance student learning rather than as a mechanism for shifting blame. Teachers cannot be responsible for the diversity of cohorts and students they teach, nor can they be expected to do their best without both professional support and resourcing from governments.

Providing effective system support is about far more than simply obtaining data. Governments must recognise their responsibilities to allow teachers, schools and education departments to put measures in place to address issues identified and the particular needs of different cohorts and students. The difficulty has never been identifying underprivileged cohorts or schools in need of assistance, but to persuade federal, State and Territory governments to prioritise these needs and supply the required resources.

The roles and responsibilities of teachers, schools, education departments and governments would be enhanced by the development of a framework for professional accountability, one that understands the roles and relationships of everyone in the education process and what their appropriate responsibilities are. Federal, State and Territory governments are presently moving towards the development of a national curriculum. Such a curriculum should include such a framework, along with a curriculum guarantee to ensure that all students, no matter where they live or what their socio-economic background is, are guaranteed access to a rigorous, rich and rewarding curriculum aimed at equipping them with the knowledge and skills required to be successful in modern society and pursue the widest range of opportunities in life.

## **Conclusion**

The present momentum towards testing has to be seen in the context of the wider education debate. Performance testing is part of the “back to basics” education movement, which argues for a greater emphasis on a rote learning approach to literacy and numeracy and a narrowing of the curriculum to allow it. Yet education in the modern world must be broadly based and give students the skills and flexibility to acquire new

---

<sup>46</sup> Australian Education Union Curriculum Policy, 2007, 5.3.3.  
<http://www.aefederal.org.au/Policy/Curric2007.pdf>

knowledge in a changing society. A narrow curriculum built around testing will not provide this. Subject and teaching methodologies that create flair and encourage thinking processes slip into the background, despite evidence that it is these flexible and creative skills that equip students best for the future, and are often what employers explicitly say they want.

The back to basics movement is accompanied by a distorted sense of educational crisis and claims modern teaching methods have led to a decline in literacy and numeracy levels in schools. It is not a claim supported by the evidence. The Longitudinal Surveys of Australian Youth (LSAY) 1975-1998 shows literacy and numeracy levels have remained constant over the last thirty years and improved for some groups, despite challenges of changing demographics.<sup>47</sup> The international PISA data show Australian students doing well by international standards. Nevertheless, claims of a crisis in literacy and numeracy have gained currency in the media and apparent acceptance by some decision makers. It is in this context that there appears to be a growing tendency to ignore past evolution and current educational expertise and re-invent curriculum based on the views of a narrow range of discipline “experts”, often from non-educational backgrounds.

There is nothing new about laments about literacy and numeracy standards; Brock demonstrates they have been a part of educational debate since time immemorial.<sup>48</sup> This does not mean that such a tendency is harmless. A false sense of crisis that is not based on evidence can be enormously damaging if it leads to the abandonment of effective teaching strategies based on the unsupported assumption that they are not working and that systemic change is necessary. This is especially so if the new strategies are as potentially harmful as are the worst excesses of mass standardised testing and its reporting. While all educational systems should clearly strive for continual improvement, an entirely different approach should be taken in the face of a system that is demonstrably failing to one for which the evidence suggests is fundamentally succeeding. The former situation would call for systemic changes, while the latter would call for assessing strengths as well as weakness and building upon them.

The latest international PISA data from 2006 suggest that Australia performs well in literacy and numeracy, ranking seventh in the OECD out of fifty-five countries for reading literacy<sup>49</sup> and thirteenth for mathematical literacy.<sup>50</sup> However, the PISA data also indicate that Australian rates poorly for equity. There is a larger gap in performance between the highest and the lowest, and a stronger than OECD average difference due to socio-economic background.<sup>51</sup> The most pressing educational issues in Australia then lie

---

<sup>47</sup> Rotham, S. 2003. *Achievement in Literacy and Numeracy by Australian 14-year-olds, 1975-1998*. Longitudinal Surveys of Australian Youth (LSAY) [http://www.acer.edu.au/documents/LSAY\\_lsay29.pdf](http://www.acer.edu.au/documents/LSAY_lsay29.pdf)

<sup>48</sup> Brock, P. “*Breaking Some of the Myths – Again*” Australian Journal of Language and Literacy, Vol. 21, 1998.

<sup>49</sup> Thomson, S. and De Bortoli, (2008) L. *Exploring Scientific Literacy: How Australia Measures Up: The PISA 2006 Survey of Students’ Scientific, Reading and Mathematical Literacy Skills*. PISA National Report. Melbourne. ACER Press. p.164.

<sup>50</sup> Ibid. p.196.

<sup>51</sup> Ibid. p.178.

not primarily with literacy or numeracy levels in themselves, or in teaching quality or methodology, but in educational equity and fairness to all.

This should not be surprising. Australia has one of the most unequal educational systems in the OECD. About a third of students attend private schools, which not only often have greater resources but are able to select students and teach to a selective cohort. These already considerable inequities increased even more under the Howard government, especially under the Socio Economic Status (SES) funding system, which has also been embraced by the Federal Labor government.

As previously mentioned, MCEETYA's own research demonstrates that public schools nationwide are under-funded to the tune of \$2.9 billion per year. Ideologically based claims that more resources are not needed, or that the equity gap can be addressed by coercive surveillance and reporting regimes are not supported by evidence nor by the consensus of educational academia. Ensuring the proper resourcing of public education must be at the heart of efforts to narrow the equity gap in education. Disadvantaged schools which cater to large numbers of students from socially disadvantaged backgrounds should not be stigmatised as "failing" schools based on simplistic performance criteria. Such schools must be provided with the additional resources they need to improve educational outcomes for their students and provide properly resourced programs established to address the needs of students with special needs. Neither can wider social inequities be ignored if educational inequities are to be addressed. In the US context Lee contrasts the failure to narrow the equity gap under NCLB with the clearly demonstrated improvement in equity in the 1970s and 1980s due to anti-poverty programs in the wider community and resourced educational programs specifically targeted at underprivileged students.<sup>52</sup>

It would be pleasing to see a widening of the education debate within Australia away from often misinformed claims about failing standards to a truly evidenced based examination of the factors that make for excellence and equity in education. Countries which consistently perform at the top of the PISA ranking such as Finland and Canada are characterised by a lack of standardised testing and league tables, high investments in public education and high salaries and status for teachers. It is puzzling indeed that Australians are being asked to emulate models from Britain and the US, countries which achieve significantly lower rankings on PISA literacy and numeracy proficiency levels than Australia, rather than countries which do better.

In order to achieve a true education revolution to meet the diverse needs of the 21<sup>st</sup> Century, Australian students must be provided with a properly resourced public education system, resources and programs specifically designed to address equity issues, and a rich, rigorous and innovative curriculum for all regardless of their background, providing them with excellence in education and the creative and critical thinking skills needed to achieve successful careers and lives in a changing society. Recognising the professionalism of teachers and schools and working with them in a supportive fashion to

---

<sup>52</sup> Lee, Op cit. p. 56.

do this is vital to achieving more successful, just and equitable educational outcomes for Australian young people.

## References

- Australia 2020 Summit. (2008) Education Skills and the Productivity Agenda. Background paper p. 10.  
<http://www.australia2020.gov.au/topics/docs/productivity.ppt#258,1> Viewed 8 April.
- Australian Education Union Curriculum Policy, 2007.  
<http://www.aeufederal.org.au/Policy/Curric2007.pdf>
- Brock, P. “*Breaking Some of the Myths – Again*” Australian Journal of Language and Literacy, Vol. 21, 1998.
- Devereaux, J. (2005) *Assessment and Reporting. A Discussion Paper*. AEU document.
- Doecke, D. Locke, T. and Petrosky, A. (2004) *Explaining Ourselves (To Ourselves): English Teachers, Professional Identity and Change*. Literacy Learning, Issue 139, pp.103-112.
- Fairtest National Center for Fair and Open Testing. *Joint Organizational Statement on No Child Left Behind (NCLB) Act*. <http://www.fairtest.org/joint-organizational-statement-no-child-left-behind> (Viewed 14 April 2008)
- Fairtest. The National Center for Fair and Open Testing. “*No Child Left Behind*” After Six Years: An Escalating Record of Failure. <http://www.fairtest.org/NCLB-After-Six-Years> Viewed 14 April 2008.
- Gillborn, D. and Youdell, D. (1998) *Raising Standards and Deepening Inequality: Selection, League Tables, and Reform in Multiethnic Secondary Schools*. Paper presented at the symposium, “Racism and Reform in the United Kingdom: The Market, Selection, and Inequality,” held at the Annual Meeting of the American Education Research Association. (San Diego, CA, April 1998). p. 12.
- Gillborn, D. and Youdell, D. (2000) *Rationing Education*. Buckingham. Open University Press.
- Gorard, S. (2006) *Value-Added is of Little Value*. Journal of Education Policy, Vol. 21, No. 2, p. 239.
- Harris, B. (2007) *Why ranking Schools Would do More Harm than Good*. Paper for the AEU Federal Executive, Melbourne.
- Kane, T.J. and Staiger, D.O. (2001) *Improving School Accountability Measures. Working Paper 8156*. National Bureau of Economic Research, Cambridge, Massachusetts.

- Lee, J. (2006) *Tracking Achievement Gaps and Assessing the Impact of NCLB on the Gaps: An In-depth Look into National and State Reading and Math Outcome Trend*. Cambridge, Massachusetts. The Civil Rights Project, Harvard University.  
[http://www.civilrightsproject.ucla.edu/research/esea/nclb\\_naep\\_lee.pdf](http://www.civilrightsproject.ucla.edu/research/esea/nclb_naep_lee.pdf)
- Linn, R.L. and Haug, C. (2002) *Stability of School Building Accountability Scores and Gains, CSE Technical Report 561*. Center for the Study of Evaluation, National Center for Research and Evaluation, Standards and Student Testing, University of California, Los Angeles.
- Luke, A. and Woods, A. (2007) “*Learning Lessons.*” *What No Child Left Behind can teach us about literacy, testing and accountability*. Australian Curriculum Studies Association. [http://www.acsainc.com.au/content/lessons\\_from\\_no\\_child\\_left\\_behind.doc](http://www.acsainc.com.au/content/lessons_from_no_child_left_behind.doc)
- Mansell, W. (2007) *Education by Numbers. The Tyranny of Testing*. London. Politicos Publishing.
- Martin, R. (2007) *Low Teacher Trust – Let’s Not Go There*. AEU document.
- National Union of Teachers. *The Case Against National Curriculum Tests*.  
[http://www.teachers.org.uk/resources/pdf/case\\_against.pdf](http://www.teachers.org.uk/resources/pdf/case_against.pdf) (Viewed 22 April 2008)
- Nichols, S.L. and Berliner, D.C. (2007) *Collateral Damage. How High-Stakes Testing Corrupts America’s Schools*. Cambridge, Massachusetts. Harvard Education Press.
- Rotham, S. (2003) *Achievement in Literacy and Numeracy by Australian 14-year-olds, 1975-1998*. Longitudinal Surveys of Australian Youth (LSAY).  
[http://www.acer.edu.au/documents/LSAY\\_lsay29.pdf](http://www.acer.edu.au/documents/LSAY_lsay29.pdf)
- Rowe, K.J. (2000) *Assessment, League Tables and School Effectiveness: Consider the Issues and ‘Let’s Get Real!’* Journal of Educational Inquiry, Vol.1 No.1. pp. 73 – 98.
- Schleicher, A. (2007) *Literacy Skills in the Information Age*. Paper presented to the 15<sup>th</sup> European Conference on Reading, Humboldt University, Berlin.
- School Performance Information. An Issues Paper for the ACT Minister for Education and Training*. (2004). Government School Education Council, ACT. 2004.  
[http://www.gsec.act.edu.au/\\_data/assets/pdf\\_file/0009/38808/School\\_Performance.pdf](http://www.gsec.act.edu.au/_data/assets/pdf_file/0009/38808/School_Performance.pdf)
- Shepherd, J. “‘Fairer’ School League Tables Misleading – Study.” The Guardian, (Manchester) January 7, 2008.  
<http://www.bris.ac.uk/Depts/CMPO/workingpapers/wp176.pdf> Viewed April 22.
- Stateline Victoria. ABC Television. 18 April, 2008. Transcript -  
<http://www.abc.net.au/stateline/vic/content/2006/s2221256.htm>

Thomson, S. and De Bortoli, (2008) L. *Exploring Scientific Literacy: How Australia Measures Up: The PISA 2006 Survey of Students' Scientific, Reading and Mathematical Literacy Skills*. PISA National Report. Melbourne. ACER Press.

Tomazin, F. "Pouring Money Into Schools 'Doesn't Work'," *The Age* (Melbourne), April 11, 2008. <http://www.theage.com.au/news/national/pouring-money-into-schools-doesnt-work/2008/04/10/1207420591367.html> (viewed 11 April)

Tomazin, F. "Gillard to Push for Parents to Get Details on School Performance," *The Age* (Melbourne), April 12, 2008. <http://www.theage.com.au/text/articles/2008/04/11/1207856836846.html>

West, H. and Hind. H. (2006) *Selectivity, Admissions and Intakes to "Comprehensive" Schools in London, England*. *Educational Studies*, Vol. 32, pp. 145 – 155.

Wilson, D. and Piebalga, A. (2008) *Accurate Performance Measure but Meaningless Ranking Exercise? An Analysis of the English School League Tables*. The Centre for Market and Public Organisation. Bristol Institute of Public Affairs, University of Bristol.- <http://www.bris.ac.uk/Depts/CMPO/workingpapers/wp176.pdf>