ISSUE ANALYSIS

THE CENTRE FOR INDEPENDENT STUDIES

No. 29 19 February 2003

The Missing Links:
Class Size, Discipline, Inclusion and Teacher Quality
A Response to the Vinson Report on Public Education in New South Wales

Jennifer Buckingham

EXECUTIVE SUMMARY

The Vinson Report on Public Education in NSW contains a wealth of information, and makes important points about schools, teachers and education in public schools throughout the state.

The Report's recommendation on class sizes has attracted more attention than any other. This is unfortunate because it is on this issue that the Report is weakest.

There are two central problems. First, the review of research findings on class size and achievement is inadequate. Second, the conclusions drawn on the basis of the limited information presented are debatable.

A more thorough appraisal of the research on class sizes reveals the following:

- Many studies have methodological problems that make their application in a real world context doubtful.
- Many studies have introduced other reforms such as curriculum changes at the same time as class size reduction, making their individual effects impossible to determine.
- The large majority of studies have found no significant effects of class size on student achievement. The remainder have shown small benefits, usually only when classes have less than 20 students.
- Reducing classes from 25 to 20 would cost an extra $1,150 per student per year, and would obtain only two minutes more individual instruction per day.
- Class size has less effect when teachers are competent.
- The single most important influence on student achievement is teacher quality.

It is far more valuable, both in educational and fiscal terms, to have good teachers than lots of teachers. The Vinson Report and the Ramsay Report of 2000 reveal that teacher education in NSW is ineffective in preparing teachers for the classroom, and that professional development is unacceptably meagre.

The Vinson Report's recommendations on class sizes will eventually cost billions of dollars and serve only to undermine the more important issue of teacher quality.

Jennifer Buckingham is Policy Analyst at The Centre for Independent Studies (CIS). She is the author of *Families, Freedom and Education: Why School Choice Makes Sense* (CIS 2001) and co-author with Ross Harold of *Issue Analysis* No.20, 'Public Education in New South Wales: Submission to the "Vinson" Inquiry' (CIS, 28 November 2001), available at www.cis.org.au

Introduction

In 2000, the NSW Teachers Federation initiated and funded an 'Independent Inquiry Into Public Education in NSW', chaired by Professor Tony Vinson. This year the Inquiry committee published its report (hereafter referred to as the 'Vinson Report').

The findings of the inquiry and the recommendations made in the three volumes of its final report received a great deal of media and political attention, and rightly so. The reports contain a wealth of information in the form of insights from students, teachers and parents, as well as previously unpublished data from the NSW Department of Education.

There are, however, two central problems. First, the committee seems to have made little effort to seek out and provide information beyond the submissions received, and only the most rudimentary of literature reviews and international comparisons are offered. Attempting to cover all the research on schooling would have made the report unwieldy and time-consuming, but there are important reasons to be thorough. Issues such as class size, where expert opinion is far from unanimous, require detailed analysis at a primary source level. Further, the majority of submissions were from teachers, who are arguably (if understandably) biased toward smaller classes.

What matters in a classroom more than anything else, including the number of students, is good teaching.

Second, the conclusions drawn on the basis of the information presented are debatable, and connections between the various troubles in schools are often not made. Anyone familiar with educational research and aware of the challenges that classroom teachers face on a daily basis knows that the difficulties associated with large classes are related to discipline problems and the wide range of abilities in each classroom. Similarly, what matters in a classroom more than anything else, including the number of students, is good teaching. The Vinson report does not make these important points explicit.

A thorough review of the research on class size and student achievement shows that much of it is flawed in ways that make it unreasonable to expect the same results in a real-world situation. Many studies have introduced other reforms at the same time as class size reduction, making the effect of class size alone impossible to determine. In most cases the persons participating in the experiment were motivated to produce positive results. Only a small minority of studies found any positive effect of smaller classes on student achievement, usually in classes of less than 20, and few of these effects were large.

The findings on class size suggest that there is little if any reason to believe that reducing classes from 25 to 20, as recommended by the Vinson Report, will have an effect large enough to warrant the cost. Research tells us that effective teaching is much more important than the number of children in the classroom. It is, therefore, much wiser to invest in the quality of teachers, rather than quantity.

Given the authority the Vinson Report and its principal author have been afforded,[1] and the likelihood that the report will be referred to regularly in the future, it is necessary to point out its flaws and put reservations with its findings on the public record.

Class size and achievement

In the area of school reform, class size reduction seems to hold all the aces. It is popular with academics, teachers, students and parents alike. It seems intuitive that to have fewer children in a class is better.

Research appears to confirm this. Several large scale studies and many smaller ones find a relationship between learning and class size. But a closer examination reveals crucial methodological problems and generalisations that make the findings less than definitive, even meaningless.

Reviewers of this research, who present it as evidence for the importance and efficacy of class size reduction, often either ignore these problems or acknowledge them in passing.

The Vinson Report had the scope and expertise to cover the issue of class size thoroughly, but it relates the findings of various studies, often from secondary sources, without the important caveats. These caveats are such that much of the research is inapplicable in other contexts. That is, the same results cannot be expected under different circumstances.

The report dismisses these problems. It concludes that the evidence of a relationship between smaller class size and better learning outcomes is strong and that the effect is large enough to warrant the expense of such reforms in NSW state schools. It recommends reducing class sizes in NSW state schools from a maximum of 25 to a maximum of 20 in Kindergarten through to Year Two (K-2).[2]

Although class size reduction is one of the most expensive reforms proposed by the Vinson Report, only seven pages are devoted to justifying it. The merits of smaller classes are considered self-evident and unarguable, yet the report's literature review is incomplete and insufficient to confidently draw these conclusions.

Hundreds of studies can be cited on the relationship between class size and student achievement. Education researchers Ronald Ehrenberg and colleagues claim that

> Most have found some evidence that smaller classes benefit students, particularly in the early grades, and especially kids at risk of being underachievers. Unfortunately, most of these studies were poorly designed. Teacher and student assignments were rarely sufficiently random; a number of studies were simply too brief or too small, and too few had independent evaluation.[3]

Other researchers, such as Eric Hanushek, go further, arguing that most of these studies are not only flawed but also fail to produce convincing evidence that class size has any significant effect on student achievement.[4] Hanushek is not without his critics and some of their points of contention are persuasive.

Most [studies] have found some evidence that smaller classes benefit students . . . unfortunately most of these studies were poorly designed.

Hanushek versus Krueger

Economists Eric Hanushek of Stanford University and Alan Krueger of Princeton University have used different methods to conduct meta-analyses of studies providing estimates of class size effects up to 1994. The debate that has taken place in recent years between these two economists is very important. It deserves closer attention than the dismissive second-hand opinion offered in the Vinson Report.[5]

Hanushek is well-known for his research demonstrating that there is no direct relationship between financial resources and school performance. He claims that only a small minority of studies show a significant positive effect of smaller classes on student achievement.

Krueger is best known for his work on Project STAR. One of the largest and most influential studies of class size reduction, its results are frequently cited as proof of the benefits of smaller classes.

In a meta-analysis of 59 studies yielding 277 estimates of the effect of class size on student achievement, Hanushek found that 14.8% of these estimates were positive and significant. That is, students in smaller classes showed significantly higher achievement than their counterparts in larger classes. The remaining estimates were either insignificant (no difference in achievement—71.9%) or negative and significant (smaller classes had lower achievement—13.4%).[6]

Krueger argues that Hanushek's method of selecting studies, extracting and counting the estimates is irrational and has produced a biased result. Krueger's main criticisms are:

- The studies from which Hanushek drew the most estimates are those which produced insignificant or negative results.
- When an insignificant or unexpected result is found by researchers, it reduces their chance of publication so they often look for disaggregated effects, separating the sample into smaller sub-samples.

- This has two consequences. First, an over-representation of insignificant and negative estimates. Second, these estimates are less powerful because the sample size is smaller.
- It is, therefore, erroneous to count each of the effect estimates from multiple-estimate studies and give them equal weight as effect estimates from single-estimate studies.

Krueger proposes three alternative methods of analysis:

1. Estimates should be given weights proportional to the number of estimates yielded in the study. For example, a single-estimate study should be counted as one, but an estimate from a study yielding four estimates should be counted as one quarter.
2. Since some studies are better designed than others, these should be given more weight in the analysis. His suggested method is citation frequency; that is, studies which are referred to more often in academic literature would be given more weight.
3. Because the smaller sub-samples in multiple-estimate studies reduce their statistical power, regression analysis should be used to estimate what the effect estimate would be if the study had yielded one estimate only.

The further a statistical analysis moves from the original data, the more room for error and the less meaningful the results.

Only the first of these is convincing. If Krueger is correct that multiple estimates from one sample are biased towards insignificance and that these results have a greater margin of error, they probably should have less weight in a meta-analysis and therefore less influence on the results.

Proposed methods 2 and 3 are problematic. Regarding the second, citation frequency is not a proven indicator of quality. It may just as easily be biased toward studies with one type of result or the other. As for the third, the further a statistical analysis moves from the original data, the more room for error and the less meaningful the results.

Hanushek counters Krueger's criticisms well:

- He argues that multiple-estimate studies provide more information than a single estimate and should not be weighted less in an analysis.
- He responds to Krueger's claim of over-representation of insignificant results from multi-estimate studies by restating Krueger's own argument that insignificant results are less likely to be published, implying that there is a bias toward positive significant results in the literature.
- He dismisses the accuracy of deriving single estimates from multiple estimates on the basis that different sub-samples of students (for example, disadvantaged students) will yield different results. This important information is lost with aggregation.

Table 1. Krueger's (2002) Re-analysis of Hanushek's (1997) Meta-analysis

	Hanushek: Estimates weighted equally	*Krueger (1):* Estimates weighted by inverse of number of estimates in study	*Krueger (2):* Estimates weighted by citation frequency	*Krueger (3):* Estimates derived from regression analyses of original estimates
Result				
Positive and significant	14.8%	25.5%	30.6%	33.5%
Insignificant	71.9%	61.2%	62.3%	58.4%
Negative and significant	13.4%	10.3%	7.1%	8.0%

Source: Lawrence Mishel and Richard Rothstein (eds), *The Class Size Debate* (Washington DC: Economic Policy Institute, 2002).

Whether one is persuaded more by the case presented by Hanushek or by Krueger, the strongest evidence is in the statistics produced by their various methods of analysis.

Table 1 shows that even when estimates are weighted and manipulated so as to avoid perceived bias toward studies showing no effect of class size—arguably creating bias in the opposite direction—the statistics do not show the 'systematic evidence of a relationship between class size and achievement' claimed by Krueger.[7]

If we accept Krueger's first and least controversial proposal—that multiple estimates from a single study should not carry as much weight as a single estimate (which is debatable even so)—only one in four studies found that students in smaller classes had achievement rates significantly higher than students in larger classes.

Only one in four studies found that students in smaller classes had achievement rates significantly higher than students in larger classes.

Other evidence

The above conclusion is consistent with the findings of other literature reviews. The Vinson report describes two national data analyses and four literature reviews as follows.

National data analyses:

- Wenglinsky (1997): In Years 4 and 8, 'lower student/teacher ratios were positively related to higher mathematics achievement'.[8] (Inconsistently, the report does not dismiss this finding due to the use of student/teacher ratio instead of class size, but does so with regard to the work of Eric Hanushek.)
- Rees and Johnson (2000): '. . . no evidence that smaller classes alone led to greater student achievement'.[9]

Literature reviews:

- Glass and Smith (1979): '. . . the major benefits of reducing class size occurred where the number of students was less than 20'.[10]
- Robinson and Wittebols (1986): 'positive effects were less likely if teachers did not change their methods and procedures in the smaller classes'.[11]
- Slavin (1990): Found that classes of less than 20 had a 'small positive effect on students that did not persist after they were removed from the smaller class'.[12]
- Hanushek (1998): 'The evidence about improvements in student achievement that can be attributed to smaller classes turns out to be meagre and unconvincing'.[13]

Of the above six studies, three conclude that there is no lasting benefit to students of reducing class size, two conclude that classes must have less than 20 students to make a difference and one found that the effect of class size was mediated by teaching style.

As well as these reviews, the Vinson Report details the findings of three major studies they describe as 'trial programmes and large field experiments'—Project STAR, SAGE and 'Prime-Time'. Each of these is presented as proof positive that smaller classes are beneficial to students. Below, the Vinson Report's comments will be summarised, followed by a more accurate representation of the studies' findings.

Project STAR (Student Teacher Achievement Ratio) in Tennessee:

According to the Vinson Report:

This is the 'most scientifically rigorous' and 'best-designed field experiment ever'.[14] The findings reported are that the positive effects of small classes (13-17 students) in K-3 on achievement levels are cumulative (the longer the time spent in a small class, the larger the effect) and persistent (the effect lasts into later grades when student return to regular size classes). It also reports that gains were greater for disadvantaged students.

The Vinson Report acknowledges that the non-random self-selection of schools into the project may be a problem, because such schools might have a greater interest and enthusiasm for such reforms, perhaps inflating the results.

Missing from the Vinson Report:
The source of the Vinson Report's information on Project STAR is not clear, but recent analyses of the Project STAR data by its principal researchers is less straightforward. In a 2001 article, Jeremy Finn and colleagues reported that the gains made by small class students on their regular class peers declined when they returned to regular classes, and that significant enduring effects of class size occurred only for students who had been in a small class for three or four years. There was only weak and mixed evidence of a larger effect for minorities.[15]

Another study from principal researchers on Project STAR found that classroom practices differed between the small classes that achieved the largest and smallest gains.[16] That is, small class benefits were mediated by the quality and method of teaching.

Although it makes a nod to it, the Vinson Report does not explain the full ramifications of the fact that Project STAR suffers from the methodological problem of the 'Hawthorne Effect'. This is where the participants in an experiment are aware of their role and the potential consequences. Caroline Hoxby[17] explains that this causes three problems: First, incentive conditions are altered, so that results produced under experimental conditions may not necessarily be the results in reality. Second, some people temporarily increase their productivity while being evaluated, especially if they have an interest in the experiment succeeding. Third, people sometimes undo the randomness of the experiment due to external pressures, for instance by placing certain children in small classes due to demands from parents.

The methodological problems of Project STAR cannot be dismissed as 'criticisms'. They create serious doubt over whether the results achieved by Project STAR would be replicated under different conditions.

The methodological problems of Project STAR cannot be dismissed as 'criticisms'.

Even if these doubts could be set aside, the findings are inconsistent with the recommendations made by the Vinson Inquiry. Small classes in Project STAR are 13-17 students. Barbara Nye of Tennessee State University, who has studied the results in detail, has been quoted as saying that 'the public shouldn't necessarily expect the same results from classes of around 20 as those of 15. It's taken a long time to get that message across'.[18] It seems the message still has a way to go.

Not only does the Vinson Report recommend that class sizes be reduced to a number that has not been shown to have any effect, it also recommends doing this in classes from Kindergarten through to Year 2 at the same time. In Project STAR, enduring results were only found for students who had been in a small class for three or four years. This suggests that it would not be effective to reduce classes in all year levels at once, but to stagger class size reduction, beginning with a kindergarten cohort.

Given that the Project STAR findings could be viewed as irrelevant, it may seem futile to point out how they should correctly be interpreted. But the fact that they have been reported inaccurately and without sufficient thought to their implications indicates the incautious approach taken to this research in the Vinson Report.

The SAGE (Student Achievement Guarantee in Education) in Wisconsin:

According to the Vinson Report:
Under the SAGE programme, K-3 classes were reduced to an average of 15 in schools where at least 50% of students were living below the poverty line. Findings cited are those from a 1999 study showing that 'Year 1 students in the SAGE program achieved better test results than students in comparison schools in language, arts and maths. Results from grades two and three generally follow the same pattern'.[19]

Missing from the Vinson Report:
Recent evidence published by the same researchers[20] confirms that students in SAGE schools performed significantly better than students in comparison schools on a variety of measures.

Most importantly, however, this cannot be attributed to reductions in class size. Schools volved in the SAGE programme implemented a variety of reforms at the same time:

1. Class-size reduction;
2. A longer school day and increased collaboration with community organisations;
3. A more rigorous academic curriculum;
4. Staff development and accountability mechanisms.

In addition, the same team of researchers discovered important differences in teaching styles between SAGE and comparison schools. Instruction in SAGE schools was predominantly teacher-centred as opposed to student-centred.[21] Differences were also identified between classrooms within SAGE schools. Teachers in higher achieving classrooms showed a preference for structured, goal-oriented, explicit instruction and classrooms with established routines where learning proceeds sequentially and at a quick pace. Teachers in lower achieving classes tended to believe that the primary advantages of reduced-size classes are the opportunities to develop critical thinking, to permit students to choose their activities and to have more activities and problem-solving lessons. They also had a more permissive management style and a more random lesson structure.

So, as in Project STAR, the aptitude of the classroom teacher is the key, not the number of children.

The aptitude of the classroom teacher is the key, not the number of children.

Prime-Time Project in Indiana:

According to the Vinson Report:
The initial results of a two year study in 24 schools where classes were reduced to an average of 18 were 'so promising'[22] that K-3 class sizes were reduced in all Indiana state schools. One analysis apparently found 'substantially larger gains in reading and maths achievement for students in small classes'.[23]

The Vinson Report gives a more accurate representation of the value of this study than it does of STAR and SAGE. It notes that the study was not random, that other changes in school policy occurred at the same time and raises the possibility that teachers were motivated to ensure that small classes worked.

Missing from the Vinson Report:
The extension of class size reduction from the original 24 schools to all schools occurred after only one year. Even reviewers who favour class size reduction have admitted it was therefore 'not possible to compare results for small classes with a comparable group of larger classes'.[24] The results cited in the Vinson Report were actually from a study of data collected before project Prime-Time was initiated.

Several other large scale studies have been conducted, the results of which are not presented in the Vinson Report. They are summarised briefly below.

California Class Size Reduction Initiative:
Inspired by Project STAR, K-3 class sizes in all Californian schools were reduced from a maximum of 33 (average 29) to a maximum of 20. To meet this requirement, schools were forced to hire underqualified teachers.

The Class-Size Reduction (CSR) Research Consortium concluded on the basis of four years of data analysis that 'no strong relationship can be inferred between achievement and CSR'.[25] Christopher Jepsen and Steven Rivkin found that the large number of extra teachers demanded by CSR led to a deterioration in teacher quality, which in some cases fully offset any benefits of smaller classes.[26]

Hoxby's (2000) Population Variation Study in Connecticut:
In this study, Caroline Hoxby,[27] a Harvard economist, looked at the relationship between achievement and changes in class size due to natural variation in age cohorts in the population. This observational approach avoids possible experimental manipulation effects. She uses two different methods to compare the class size and achievement of adjacent cohorts, taking into account enrolment data and maximum class size regulations.

Neither method shows that smaller classes produced achievement gains. Even given the precision of the data analysis, which allowed tiny improvements to be significant at the 5% level (the improvements found in Project STAR would have been significant if found in this study), the effect of reducing class size was estimated to be close to zero. Further, the results do not suggest that class size reductions are more effective in schools that serve low-income or African American students (in fact, the only significant result was an improvement in fourth grade reading scores of high-income students).

Investment in fewer, more highly educated and better paid teachers seems to result in higher student achievement.

Christchurch Health and Development Survey:
A longitudinal study conducted in New Zealand, although not designed to study class size effects, has yielded information that can be used as an observational study.

Michael Boozer and Tim Maloney[28] first compared the results of children permanently in small (19), medium (29.9) and large (33.8) classes between the ages of 8 to 13 years. Only a small number of students were permanently in classes of these sizes over the age period, and the results were insignificant. They then compared students whose *average* class size over this age period was small (21.2), medium (29.7) or large (33.2). They found significant effects only for children in *persistently* smaller average classes between the ages of 8 and 13, on both childhood test score improvements as well as on early adult outcomes such as completed education and unemployment.

UK National Child Development Study
In another observational study of existing data from the 1960s, Maria Iacovou[29] controlled for school type/size and streaming to account for the possibility (and some evidence) that less able children are more likely to be allocated to a smaller class—which would make the variation in achievement in different size classes internally created.

Iacovou looked at average class size at age 7 (excluding students in classes of less than 20 and more than 45) and found that class size was related to student attainment in reading but not maths. A smaller effect persisted to age 11 only for girls and for children from large families. There was no evidence of greater benefit to disadvantaged students.

Third International Maths and Science Survey (TIMSS)
Class size effects for 18 countries were estimated using maths and science performance in TIMSS and average class size data. Ludger Woessmann and Martin West[30] found that class size effects varied greatly between countries, with large effects in only two countries: Greece and Iceland.

When they compared these countries with those where no class size effect was found, several things were apparent. First, countries with large class size effects performed below average internationally, whereas those with small or no class size effects performed above average internationally. Also, countries with large class size effects had less educated, lower paid teachers compared to countries with small or no class size effects.

From this they drew several conclusions. First, class size effects cannot be imputed from one country to another because school systems vary significantly. Second, class size is more important when teachers are less effective. Investment in fewer, more highly educated and better paid teachers seems to result in higher student achievement.

Australian research

Australian research on class sizes is scarce. A study by Sid Bourke in Melbourne in the 1980s found that smaller classes were related to higher achievement in maths,[31] but John Keeves has noted that analysis of these results at the class level revealed that class size was also related to student ability (sorting) and that controlling for this changed the relationship between class size and achievement. Keeves concludes that 'there is little clear evidence to support the costly reductions in class size'.[32]

The analysis of TIMSS results described earlier did not lead to any meaningful findings for Australia. The researchers found that average Australian class sizes in maths and science were not good proxies for actual class sizes, so differences in student achievement between classes of different sizes could not be confidently attributed to the size of the class.

Implications

The Vinson Report estimates that the reduction of class sizes to a maximum of 20 in Years K-2 would cost $47 million dollars per annum in disadvantaged schools and $225 million per annum in all schools. This is the most expensive recommendation made, the all schools figure of $225 million exceeding the total cost of all other recommendations by 40%.

Even this figure underestimates the cost of class size reduction as it accounts only for extra staffing costs. Each additional teacher necessitates an additional classroom, must be educated and trained, will need extra classroom resources and require on-going professional development. The cost of more classrooms has been conservatively estimated by the NSW Opposition to be in the order of $140 million initially.[33]

Not only is the cost large, but the findings of the studies described above are mixed and weak at best on the issue of class size. Only one thing comes through loud and clear from the research: what goes on in the classroom is more important than how many children are involved. This is not to say that classroom activity is unaffected by the number of children, but that proven and appropriate teaching methods are paramount.

What then does the Vinson Report make of this? It recommends that large scale class size reduction takes place in state schools, bringing class sizes in K-2 to a maximum of 20. The report says that it has been guided 'not only by the consistency of the findings, but also the quality of the research yielding particular results'.[34]

Much of the Vinson Report's information on class size research comes from a short literature review by Bruce Biddle and David Berliner, including their conclusions, which are reproduced verbatim. Yet Biddle and Berliner seem just as confused as the authors of the Vinson Report, claiming that 'although the results of individual studies are always questionable, a host of different studies suggest several conclusions',[35] namely that class size reduction is beneficial for students in the short and long term in academic achievement and other outcomes. In other words, these authors seem to be saying that a large number of poorly designed studies with mediocre results can be amassed into strong evidence of a significant effect.

Reducing class size by the amount [the Vinson Report] recommends would not justify the expense.

Even less convincing is the Vinson Report's attempt to justify their recommendation in the face of the evidence they have presented to the contrary. They argue that policymakers should not 'await an unlikely total consensus . . . but to base policy on the best available information, after considering the strengths and limitations of the research'.[36] Complete agreement from researchers may be too much to ask, but if the Report's authors follow their own advice and seriously consider the evidence presented, notwithstanding the evidence they neglected, they would have to conclude that the best available information is that reducing class size by the amount they recommend would not justify the expense.

Theories and fallacies of class size reduction

There are several theories as to why smaller classes should be beneficial:

1. Increased individual attention and instruction;
2. Greater scope for innovation and student-centred teaching;
3. Increased teacher morale;
4. Fewer disruptions.

The idea that a teacher can devote more time to each student in a smaller class, thereby increasing the amount students learn, is the most intuitively appealing of all these theories. Yet simple calculations show this appeal to be misplaced.

In a six hour school day, approximately five hours is spent in the classroom. If half this time is spent directly addressing the class, and the other half on individual attention, each child would hypothetically receive six minutes of individual instruction in a class of 25 and 7.5 minutes of individual instruction in a class of 20. That is, an extra $1150 per student per annum[37] buys an extra 1.5 minutes per day of teacher's time. If two-thirds of classroom time is spent on individual attention, students get two minutes more in a class of 20 than 25.

These calculations may be simplistic, but indicate the insubstantial change in individual attention that a 20% reduction in class size brings, at considerable cost.

Another counter to the individual instruction theory comes from Project STAR. Some of the regular size classes were assigned a teacher's aide. Even though children in these classes presumably had twice as much individual attention, there was no difference in achievement levels between regular size classes with and without teacher's aides.

Reducing class sizes will have little or no effect without ensuring that teachers adopt instruction and management practices proven to be effective in small classes.

The second theory—that small classes provide the potential for more effective teaching strategies—suggests that class size may be conducive to greater student achievement but does not guarantee it. It also suggests that small classes alone do not produce gains in learning; that their benefits are mediated by teacher quality. Research discussed earlier demonstrates that there were notable differences in teaching and classroom management styles between high and low achieving small classes.

Teachers rarely change their teaching and classroom management styles. Even Project STAR data shows this, with few teachers modifying their classroom practices in different size classes after attending a professional development programme.[38] If this is the case, then reducing class size will have little or no effect without ensuring that teachers adopt instruction and management practices proven to be effective in small classes. This substantial investment in professional development once again adds to the cost of class size reduction, and would more than likely be equally effective without changing class sizes.

The last two theories of small class benefits are related and are the most convincing. Small classes are overwhelmingly popular with classroom teachers and it is not difficult to understand why. Schools are being forced to cope with, and attempt to educate, an increasing number of students who are uninterested and badly behaved. In some areas of Sydney, schools have difficulty attracting and retaining teachers primarily for this reason and teachers in all areas are finding their jobs more and more difficult and stressful.

Fewer students like this in a class would make teaching much easier. Reducing class sizes might be justifiable if it can be shown that the increased cost of reducing class size is offset by the decreased cost of teacher attrition, stress and sick leave.

It must be ensured, however, that a new demand for teachers does not result in the same situation as in California, where the least qualified and least experienced teachers were disproportionately employed in the most disadvantaged schools. The most simple and effective way to avoid this is to offer financial incentives for teachers in difficult-to-staff schools, which means departing from rigid wage structures based on years of service.

The impact of discipline and inclusion policies

Commonsense says that, within wide limits, class size makes very little difference when you have well-behaved students of similar ability. The number of students becomes more significant for a class of mixed ability with one or two students with disabilities and a few disruptive students. Over the past decade at least, there has been an increase in both disciplinary problems and in the number of students in state schools with disabilities of various kinds. It must be acknowledged that the strain of large classes is strongly related to class composition.

Data from the Annual Reports of the NSW Department of Education and Training (DET) show a large increase (almost double) in short and long suspensions and a small increase in expulsions from NSW state schools over the last decade.[39]

Table 2. Average Daily Rate of Suspensions and Expulsions in NSW State Schools, 1994-2001

	Short Suspensions	*Long Suspensions*	*Exclusions**	*Expulsions***
1994	108.5	16.3	1.2	
1995	127.6	18.5	1.4	
1996	138	20	1.2	1.4
1997	160	24.5	1.9	1.8
1998	176.1	26.	1.2	1.6
1999	158	24		1.8
2000	181	29		1.7
2001	207	34		1.7

* Student excluded from one state school but not all state schools. Only the Minister can expel students from the public school system.
** Called 'position declared vacant' in 1996, 1997, 1998.
Source: NSW CCG Performance Report 1997; NSW DSE Annual Reports 1996-1997; NSW DET Annual Reports 1998-2001.

The trend shown in Table 2 above is quantitative only, and any change in the seriousness of the incidents involved cannot be inferred. It is also not apparent whether the increase is due to more students being suspended or the same proportion of students being suspended more often.

As the Vinson Report demonstrates, the major problem for teachers is 'the ongoing, cumulative nature of minor infractions and disruptions, rather than their magnitude'. If these are increasing at the same rate as offences requiring suspensions, it is surprising that not more teachers have left the profession.

Efforts to downplay this problem, perhaps to protect the reputation of state schools, are counterproductive. The Vinson Report takes great pains to point out that the proportion of extremely difficult students is small (5%)[40], and the NSW DET Annual Reports make similar claims. Yet the experience and testimony of teachers and students provided in the Vinson Report reveal that bad behaviour is 'the single biggest factor that could destroy public education'.[41]

Bad [student] behaviour is 'the single biggest factor that could destroy public education'.

Teachers now have more problems with bad behaviour, but less sovereignty and authority over discipline in the classroom. They are also expected to meet the needs of students with a wider range of abilities in the one class.

Standard practice in the past has been to place students with disabilities requiring special attention in 'support classes'. Teachers of these classes were specially trained to understand the different obstacles to learning these children have and how to help them. The number of students with disabilities in NSW schools has increased markedly in the last decade or so—from 14,488 in 1988 to 35,256 in 2002—and the majority of these are now in mainstream classes. More than half of integrated students are classed as needing moderate to high learning support.[42] Such inclusion policies have contributed to the difficulties facing classroom teachers today, and the growing demands for smaller class sizes.

The Vinson Report's treatment of the topic of inclusion is lengthy, describing the particular difficulties that 'mainstreaming' presents, and ultimately finding in its favour for social reasons rather than educational. The need to evaluate the benefits of inclusion for children with disabilities in mainstream classes against the welfare of their classmates is recognised, but although research showing the benefits of inclusion to children with disabilities is cited, no evidence of the effect on other children in the class is presented.

The Report acknowledges the improbability that a child with special needs will get the same specialised attention in a mainstream classroom as they would in a

support class with a properly trained teacher. Teachers' concerns that they are unequipped to provide disabled students with the help they require, especially in secondary school, are noted. It is also apparent that inclusion has not resulted in, or from, a reduction in funding. The 100% increase in students with disabilities has been accompanied by a 400% increase in funding support.

Despite the dubious educational merits of inclusion, and its high costs, the Vinson Report recommends that it be encouraged and facilitated and that all classroom teachers be trained in instructing children with special educational needs. The Report's authors' explanation is that to segregate students according to ability is 'no longer considered acceptable'. They do not attempt to argue that segregation or inclusion is 'more educationally efficient or effective' but base their claim on 'a set of values concerning what is right'.[43]

Schools do have a socialising role, but this should not take precedence over a sound education. This is not to say that schools should be completely task-oriented, with no concern for the welfare of students, but rather that educational imperatives should prevail where there is no evidence that this will harm students.

Teacher quality

Commonsense also says that it is better to have a great teacher in front of a large class than a mediocre teacher in front of a small one.

Writing in the Bulletin of the US National Association of Secondary School Principals, Leslie Kaplan and William Owings state that 'research affirms that teaching quality is the single most important factor influencing student achievement',[44] and cite a wide variety of supporting studies. Australian research has also shown that the largest differences in achievement between students is that between students in different classes. [45]

The 'Ramsay Report' on the Review of Teacher Education in NSW,[46] provides plenty of evidence to support the primacy of teacher quality, demonstrating the impact of teachers on student achievement and the benefits from investing in teacher education.

Although much has been said about the importance of teacher quality, what makes a good teacher is yet to be adequately defined. We know that some teachers bring about higher levels of achievement from their students than others, but consensus on how is still elusive.

A certain proportion of good teaching comprises temperament, charisma, enthusiasm and other qualities that cannot be measured or taught. However, several criteria can be identified:

1. Mastery of subject matter and curriculum content;
2. Awareness of the individual abilities and capabilities of students;
3. Classroom management skills;
4. Use of teaching strategies that are proven effective;
5. Good verbal communication skills.

Schools do have a socialising role, but this should not take precedence over a sound education.

Each of these capacities is necessary but insufficient on its own. Strong content knowledge is crucial but not enough—teaching also requires a set of professional skills separate from but related to the subject being taught.[47] These skills are supposed to be gained from teacher education courses.

What constitutes effective pedagogy is beyond the scope of this paper, but there seems to be agreement that teacher education in Australian universities is inadequate in imparting both pedagogical and behaviour management skills to teachers. There is too much emphasis on the theoretical over the practical—too much Bloom and not enough classroom. New teachers have usually spent only a few weeks in teaching practicum, and support for them in the extremely difficult first year in a school is patently inadequate.[48]

Another problem is the lack of ongoing professional development for classroom teachers. The NSW Department of Education's undervalues the need for teachers to

be aware of new developments in both curriculum and pedagogy, and teachers have too few incentives to seek out professional development opportunities for themselves.

Conclusions

The Vinson Report has some important things to say. The state of disrepair and lamentable lack of facilities in many public schools is unacceptable and needs to be rectified. Welfare and discipline are major issues for public schools and require urgent action. Teacher salaries are too rigid and do not reward good teaching.

Regrettably, the report is let down by a lack of rigour and an apparent partiality in its analysis of reforms that have important implications, such as class size. Evidence both presented in and absent from the report shows that most research is either flawed or shows a marginal effect of reducing class size or both. Good teaching practices are far more effective. Yet the report recommends a class size reduction strategy that eclipses all other education spending.

When it comes to teachers, quality is far more important than quantity. The recommendations on class size reduction serve only to weaken the case for more urgent and supportable interventions, such as improved teacher education and professional development.

The influence of the Vinson Report and its chief sponsor—the NSW Teachers Federation—should not be underestimated. The NSW government has responded with caution to the recommendations on class size, promising only a pilot study. The NSW Opposition has been less circumspect, making a commitment to reducing kindergarten class sizes to an average of 21 from an average of 26 if elected.

Given that good Australian data on class size effects is non-existent, and that research from other countries is inconclusive on whether there are even marginal benefits from class size reduction, it is prudent that NSW Governments seek more evidence before embarking on what will eventually be a multibillion dollar spending spree.

If, however, class size is the only variable studied, it almost certainly will be a pointless undertaking. Any study conducted must be carefully designed and implemented and must measure the effect on student achievement of both class size and teaching practices, as well as the interaction between them. Otherwise we will be none the wiser.

It is prudent that the NSW Government seek more evidence before embarking on what will eventually be a multibillion dollar spending spree.

Endnotes

1 For example, the NSW Opposition's 'Education Policy Initiative' on class sizes draws almost exclusively on the Vinson Report for justification. See www.barryofarrell.com/misc/Getting_the_Best_Start_policy.pdf

2 Report 1, p.85, Recommendation 3.3.

3 Ronald G. Ehrenberg , Dominic J. Brewer, Adam Gamoran and J. Douglas Willms, 'Does Class Size Matter?', *Scientific American* 285:5 (2001), 78-85.

4 Eric Hanushek, 'The Evidence on Class Size', *Occasional Paper* 98-1 (W. Allen Wallis Institute of Political Economy: University of Rochester, 1998).

5 Report 1, p. 84 provides a comment on Hanushek's work derived from Biddle and Berliner 2002 (see 24 for reference).

6 Percentages do not add to 100 because of rounding.

7 Alan B. Krueger, 'Understanding the Magnitude and Effect of Class Size on Student Achievement' in *The Class Size Debate*, ed. Lawrence Mishel and Richard Rothstein (Washington, D.C: Economic Policy Institute, 2002), 31.

8 Report 1, p. 83.

9 As above, n.8.

10 As above, n.8.

11 Report 1, p. 84.

12 As above, n.11.

13 Report 1, p. 84, quoted from Hanushek 1998.

[14] Report 1, p. 82.

[15] Jeremy Finn, Susan B. Gerber, Charles M. Achilles, and Jayne Boyd-Zaharias, 'The Enduring Effects of Small Classes, *Teachers College Record* 103:2 (2001), 145-183.

[16] Jayne Boyd-Zaharias and Helen Pate-Bain, 'Early and New Findings From Tennessee's Project STAR', in *How Small Classes Help Teachers Do Their Best*, ed. M.C. Wang and J.D. Finn (Philadelphia: Temple University Center for Research in Human Development and Education and the U.S. Department of Education, 2000), 65-97.

[17] Caroline M. Hoxby, 'The Effects of Class Size on Student Achievement: New Evidence From Population Variation', *The Quarterly Journal of Economics* 115:4 (2000), 1239-1285.

[18] Linda Jacobson, 'Research: Sizing Up Small Classes', *Education Week* (28 February 2001).

[19] Report 1, p.83 refers to a 1999 evaluation by Molnar, Smith and Zahorik, but a complete reference is not provided.

[20] Alex Molnar, Philip Smith, John Zahorik, Anke Halbach, Karen Ehrle, Lawrence M. Hoffman and Beverley Cross, *2000-2001 Evaluation Results of the Student Achievement Guarantee in Education (SAGE) Program* (University of Wisconsin-Milwaukee: Centre for Education Research, Analysis and Evaluation, 2001).

[21] As above, n.20.

[22] Report 1, p. 83.

[23] Report 1, p. 83, results cited are those from McGivern, Gilman and Tillitski (1989).

[24] Bruce J. Biddle and David C. Berliner, 'What Research Says About Small Classes and Their Effects', *Policy Perspectives* (San Francisco: WestEd, 2002), 6. Apart from the fact that they play down the problems associated with class size research showing a positive effect and play up the problems associated with class size research showing no effect, Biddle and Berliner make it clear where their allegiances lie through their choice of words. Critics of negative class size research 'point out', while critics of positive class size research 'pounce' and 'decry'.

[25] Brian M. Stecher and George W. Bohrnstedt (eds), *Class Size Reduction in California: Summary of Findings from 1999-2000 and 2000-01* (CSR Research Consortium, California Department of Education, 2002), 2.

[26] Christopher Jepsen and Steven Rivken, 'What is the Tradeoff Between Smaller Classes and Teacher Quality?', *Working Paper* 9205 (Cambridge, MA: National Bureau of Economic Research, 2002).

[27] Boyd-Zaharias and Pate-Bain, 'Early and New Findings From Tennessee's Project STAR', see n.16.

[28] Michael A.Boozer and Tim Maloney 2001, 'The Effects of Class Size on the Long-Run Growth in Reading Abilities and Early Adult Outcomes in the Christchurch Health and Development Study', *Working Paper* 01/14 (New Zealand Treasury, 2001).

[29] Maria Iacovou, 'Class Size in the Early Years: Is Smaller Really Better?' (Essex University: Institute for Social and Economic Research, 2001).

[30] Ludger Woessmann and Martin R. West, *Class Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS* (Research Paper PEPG/02-02, Program on Educational Policy and Governance, 2002), http://www.ksg.harvard.edu/pepg/pdf/PEPG02-02.pdf

[31] Sid F. Bourke, 'How Small is Better: Some Relationships Between Class-size, Teaching Practices, and Student Achievement', *American Educational Research Journal* 23:4 (1986), 558-571.

[32] John P. Keeves, 'The Contribution of IEA Research to Australian Education', in *Reflections on Educational Achievement: Papers in Honour of T. Neville Postlethwaite*, ed. Wilfried Bos and Rainer H. Lehmann (New York: Waxmann, 1995), 148.

[33] www.barryofarrell.com/misc/Getting_the_Best_Start_policy.pdf

[34] Report 1, p. 84.

[35] Biddle and Berliner, 'What Research Says About Small Classes and Their Effects', see n. 24.

[36] Report 1, p.81.

[37] $225 million divided by number of students (195,906); Report 1, p. 86, Table 3.1.

[38] Ronald G. Ehrenberg, Dominic J. Brewer, Adam Gamoran and J. Douglas Willms (2001), 'Class Size and Student Achievement', *Psychological Science in the Public Interest* 2: 1 (2001), 1-30.

[39] The 2000 NSW DET Annual Report claims that the increase in short and long suspensions is a result of the new long suspension and expulsion procedures introduced in 1999 and 'the increased

authority of school principals to either suspend or expel students from schools'. In this case, we would expect to see an increase in suspensions and expulsions between 1998 and 1999, but instead a decline in suspensions occurred in 1999. Moreover, procedural changes in 1999 do not explain the increase in suspensions up to 1998 and the continued increase in 2000 and 2001.

[40] Report 2, p.52.

[41] Report 2, p.53.

[42] Report 3, p.43.

[43] Report 3, p.48.

[44] Leslie S. Kaplan and William A. Owings, 'The Politics of Teacher Quality: Implications for Principals', *National Association of Secondary School Principals Bulletin* 86: 633 (2002), 22-41.

[45] Ken Rowe, 'The Importance of Teacher Quality', *Issue Analysis* 22 (Sydney: The Centre for Independent Studies, 2001).

[46] Gregor Ramsay, 'Quality Matters: Revitalising Teaching: Critical Times, Critical Choices', *Report of the Review of Teacher Education* (Sydney: NSW Department of Education & Training, 2000).

[47] Linda Darling-Hammond, 'Teacher Quality and Student Achievement: A Review of State Policy Evidence', *Education Policy Analysis Archives* 8 : 1 (2000), http://epaa.asu.edu/epaa/v8n1;

Kati Haycock, 'Good Teaching Matters', *Thinking K-16* 3:2 (1998). Susan S. Goldsmith, 'The Pedagogy of the Subject and Professional Development', in *A Consumer's Guide to Teacher Quality: Opportunity and Challenge in the No Child Left Behind Act of 2001* (Washington D.C: National Council on Teacher Quality, 2002).

[48] Ramsay, 'Quality Matters', see n.46; and Vinson Report 3, Chapter 11.

Previous papers in The Centre for Independent Studies' *Issue Analysis* series include:

No. 28. Poor Laws (2): The Minimum Wage and Unemployment *Kayoko Tsumori*

No. 27. Getting It Right Some of the Time: An Appraisal of the Report on the Inquiry Into the Education of Boys *Jennifer Buckingham*

No. 26. Poor Laws (1): The Unfair Dismissal Laws and Long-term Unemployment *Kayoko Tsumori*

No. 25. Whose Progress?: A Response to the ABS Report *Measuring Australia's Progress* *Peter Saunders*

No. 24. Families, Fertility and Maternity Leave *Barry Maley*

No. 23. Poor Statistics: Getting the Facts Right About Poverty in Australia *Peter Saunders*

No. 22. The Importance of Teacher Quality *Ken Rowe*

No. 21. Poor Arguments: A Response to the Smith Family Report on Poverty in Australia *Kayoko Tsumori, Peter Saunders and Helen Hughes*

No. 20. Public Education in New South Wales: Submission to the 'Vinson' Inquiry *Ross Harold and Jennifer Buckingham*

No. 19. Towards Racial Harmony: A New Constitution for Fiji? *Wolfgang Kasper*

No. 18. Breaking the Trade Stalemate: What are Australia's Options? *Wolfgang Kasper*

No. 17. School Funding for All: Making Sense of the Debate over Dollars *Jennifer Buckingham*

No. 16. Code of Silence: Public Reporting of School Performance *Alison Rich*

No. 15. Setting the Record Straight: Free Trade and the WTO *David Robertson*

No. 14. Playing with Fire: Churches, Welfare Services and Government Contracts *Samuel Gregg*

No. 13. The Truth About Private Schools in Australia *Jennifer Buckingham*

No. 12. Degrees of Difficulty: The Labour Market Problems of Arts and Social Science Graduates *Andrew Norton*

No. 11. Trailing the Class: Sole Parent Families and Educational Disadvantage *Alison Rich*

No. 10. Shooting the Messenger: A Critique of Australia's Internet Content Regulation Regime *Heath Gibson*

No. 9. The Puzzle of Boys' Educational Decline: A Review of the Evidence *Jennifer Buckingham*

No. 8. Noble Ends, Flawed Means: The Case Against Debt Forgiveness *Ian Harper, Helen Hughes, Samuel Gregg*

No. 7. Taxi! Reinvigorating Competition in the Taxi Market *Jason Soon*

No. 6. Why Small Business is Not Hiring: Regulatory Impediments to Small Business Growth *Jason Soon and Helen Hughes*

No. 5. Rear Vision on Trade Policy: Wrong Way, Go Forward *Ron Duncan*

No. 4. The Last Refuge Hard and Soft Hansonism in Contemporary Australian Politics *Chandran Kukathas and William Maley*

No. 3. Tax Injustice: Keeping the Family Cap-in-Hand *Lucy Sullivan*

No. 2. Equal Pay for Work of Equal Value: Moving Toward, or Away From, Wage Justice for Women? *Helen Hughes*

No. 1. Open for Business? Australian Interests and the OECD's Multilateral Agreement on Investment (MAI) *Wolfgang Kasper*

Publications in the *Issue Analysis* series are subject to a reviewing process.

THE CENTRE FOR INDEPENDENT STUDIES