Selecting reference schools to estimate a national resource standard:

which measures of student outcome?

Richard Teese

Prepared for the:

Department of Education, Employment and Workplace Relations

October 2011

THE UNIVERSITY OF MELBOURNE

CENTRE FOR RESEARCH ON EDUCATION SYSTEMS

MELBOURNE GRADUATE SCHOOL OF EDUCATION

Table of Contents

List of Tables and Charts

Charts

Tables

The aim of this analysis is to examine the suitability of measures to guide the selection of "reference schools" for the purpose of estimating a national resource standard for Australian schools. Besides technical and practical issues, there are broader issues which relate to choosing a good set of reference schools. We begin with a discussion of four of these issues before turning to the measures themselves.

Broad issues
1 Variable cost factors
2 Selecting schools using student outcome measures
3 Whole-of-cohort measures
4 Purposive sampling of characteristic settings

Variable cost factors

To estimate a resource standard, reference schools are required which between them display a range of variable cost factors. Six such factors are briefly discussed below:

Variable cost factors to be reflected in the validation study of reference schools
1 Student numbers
2 Teaching costs
3 Student characteristics and needs
4 Parental preferences
5 School programs
6 Student outcomes

(1) *Variable student numbers*. Small schools are costly to run because minimum running costs must be met regardless of enrolments. In remote and very remote areas, the costs of provision are high, but it is less often recognized that there are also small schools in metropolitan and provincial centres and that their costs, too, are high. From one angle, this represents an inefficiency because it should be avoidable. But metropolitan and provincial schools may become small due to the way resource arrangements work and to local "divisions of labour" as to which students are educated where, Some schools become "refuge

schools", not because they are inefficient or ineffective, but because they educate a particular spectrum of the local community and are, in effect, used by other schools to offload or avoid more challenging students. Some schools, on the other hand, are small because they serve particular communities of faith who elect to educate their children in separate establishments. The running costs of these schools will also be high. At the other end of the size range are very large schools with access to scale economies.

(2) *Teaching costs*. The cost of teaching staff in schools of the same size can vary widely as a function, e.g., of levels of appointment and responsibilities. Cost is a different dimension to staffing numbers (FTE allocation): two schools may have the same allocation or complement of staff, but quite different cost profiles. In the research literature, the impact of teacher cost differences is debated. But in the market place, it is not. Well-educated parents expect schools to have highly qualified and experienced staff, no doubt with newer, less experienced staff at different stages in their professional development. Schools that struggle to maintain or build a high quality staffing profile also struggle to get good results.

(3) *The characteristics and needs of students.* Addressing the needs of children in schools serving disadvantaged communities imposes higher resource requirements on schools, not only in respect of *numbers* of staff, but as regards the depth of their training and experience and the support they themselves require. As the quotation in the box below makes clear, student needs are complex and generate extra demands on teachers, especially when there is a high concentration of disadvantage in a school and multiple forms of disadvantage.

> "*We have had to implement programs that are aimed at the 0-4 age bracket and parent education as well as fund additional expertise in the area of literacy and numeracy to support student progress. Attendance at school also impacts on student learning, particularly Indigenous students. Many students come to school having been exposed to high levels of violence as a result of domestic violence, alcohol and drug issues. These children can present as highly traumatised, and behaviour then impacts on their learning and greatly affects the teaching and learning programs as a considerable proportion of the school day is taken up with managing student behaviour*"
>
> *(Western Australian primary school principal)*

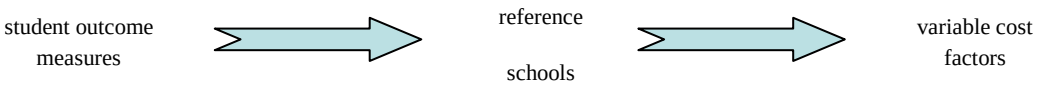
(4) *Parental preferences*. Parents may have high expectations regarding the range of programs, activities and specialist support services that a school should offer. In schools serving more advantaged communities, parents typically expect a high level of individual care and supervision and a high level of academic success.

(5) *Program range*. A fifth variable cost factor relates to the range of programs and activities which schools provide. The pressure to differentiate curriculum is strong in schools serving poorer communities, e.g., to offer nationally accredited vocational certificates. Creating student access to these programs involves a variable cost (e.g., through purchasing arrangements). But ensuring that all key areas of the academic curriculum are staffed by qualified teachers may also impose higher costs (e.g., salary premiums to attract or retain specialist staff). However, range of programs is not only an issue at upper secondary level, where it is more obvious. Primary schools also need to be able to offer a breadth of activities and programs to ensure that they engage all children effectively and that the whole learning potential of a child is developed. Areas such as music and languages are typically hard to offer within resource constraints and are variable in provision. Given widening achievement gaps, program differentiation such as through the creation of specialist remedial classes is vital in many settings.

(6) *Student outcomes*. While NAPLAN tells us about national literacy and numeracy standards, schools work on a far wider front (as expressed in the Melbourne Declaration). A resource standard is not only about "the basics". Greater depth of learning and greater breadth in the social range of successful learners are two major goals of Australian schooling. A resource standard has to address these broad requirements. In some secondary schools, for example, there is pressure to ensure that every student reaches university; in others it is to see that most complete school and that many find an apprenticeship or a good job; in others that all students have strong personal and organizational skills, high self-esteem, and sound basic skills. Depending on the context, these different outcome profiles impose different sets of costs.

Selecting schools using student outcome measures

To estimate a national resource standard for Australian schools on the basis of the costs of quality schooling in a set of reference schools assumes that the sample of schools will reflect the range of cost pressures to which schools in their diversity are exposed. If the selection of the reference schools is to be determined by measures of student outcome, these measures themselves must be broad enough to capture the different cost pressures experienced by schools that are successful in different circumstances.



To illustrate the cost sensitivity of student outcome measures, take the example of the 80%+ criterion recommended by the ACG report. This criterion is formulated as 80% + of students above national minimum standards in NAPLAN. If applied to Western Australian government primary schools, this criterion would select schools which rank at half of one standard deviation above the ICSEA mean.

Figure 1 WA government primary and district high schools: mean SES by percentage of Year 3 students above national minimum reading standards

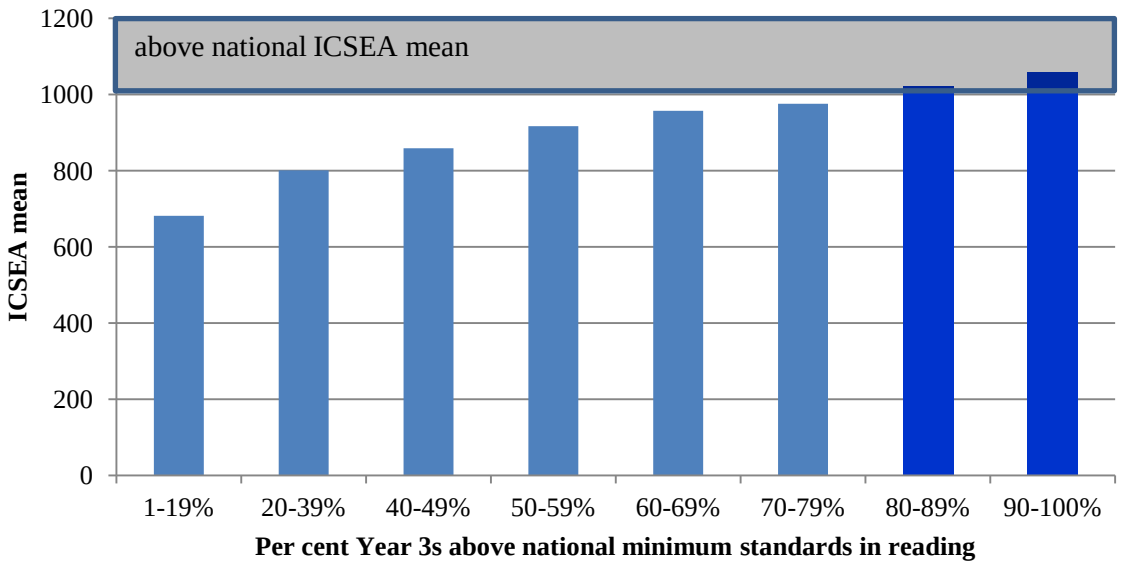


Figure 1 shows that each increment in performance as measured by percentage of Year 3 students above national minimum standards is accompanied by an increment in SES. The 80%+ criterion also picks many non-government schools, including those operating at high

levels of resource use. Socially the group of high performing schools is biassed upwards, although it is also diverse. Small remote or very remote schools enter the picture because they, too, perform very well (though not necessarily consistently, which is part of the ACG criterion). Capturing schools which serve mainly high SES families is important, but the balance of cost pressures in these schools is different to the balance in poorer schools. For example, parental demand for a high level of program enrichment and ancillary services will be high in these schools and drive up costs, producing a different factor profile than in the case of small rural schools which are also costly to run.

One question is whether the outcome criterion will gain access to schools with the cost profile more typical of schools serving poor communities. Of high performing schools that reach the performance threshold, but are *below the national ICSEA mean*, many are small provincial or remote/very remote schools (representing 40% of this group), while 60% are metropolitan. A more detailed analysis than is possible here would be required to examine the characteristics of the schools that are low SES, but meet the performance threshold on NAPLAN proposed by the ACG report.

Whole-of-cohort measures

High quality schooling is an entitlement of all students, and measures of outcome should reflect the experience of all students. Examples of whole-of-cohort measures are standardized tests, student attendance, and student satisfaction. A measure that approaches whole-of-cohort status is school completion—though students leaving from below Year 12 to take up apprenticeships are not reflected in this measure. Exam results are only a partial measure of cohort success. This is because not all senior students undertake graded assessment of their work (e.g., VCAL in Victoria), and legally are not required to do so (e.g., the VCE). There are other problems with exam results, but the focus in this section of the discussion is on the need to measure quality as experienced by the whole cohort. An example of a partial measure is ATAR (discussed in detail later). Finally destination measures have the potential to reflect the experience of the whole cohort, but this depends on design, methodology and survey response rates (where relevant).

Schools operate policies which are not always inclusive in impact. For example, secondary schools may not offer much in the way of VET options or other alternatives which assist weaker students to better engage with learning. Such schools may discourage more difficult students from enrolling or staying on, and may seek to specialize in academic subjects, leaving other schools to manage rejected students. Success of a certain kind may follow from this philosophy, e.g., better overall exam results or greater success in rankings for university selection. But this shifts the costs of inclusiveness policies to other schools, and these are unlikely to perform as well academically had they been more selective and less accepting of the behaviour of neighbouring schools.

Measures that capture the success of only part of the cohort will not reflect the costs of provision for other groups. For example, in a secondary school where only a minority of students apply for university, there are costs associated with the delivery of VET.

Purposive sampling of characteristic settings

The search for universal measures of school performance may not be the most productive approach to choosing reference schools for a validation study. While the availability of measures, such as NAPLAN, enables most schools to be ranked and high performing schools to be quickly identified, there are also weaknesses with this approach (discussed under the heading of each measure). A purposive sampling strategy also has advantages. Most importantly, it gives access to measures which at present are available only in some jurisdictions (and only partially within these). An example is post-school destinations. Secondly, a purposive sample can be constructed explicitly around known variable cost factors which are of great relevance to setting a resource standard. Thirdly, choice of a range of different contexts can be carefully managed and verified as representative. Fourthly, student outcome measures can still be used to capture high quality schooling, but with less risk of social or academic bias and with more focus on differential costs.

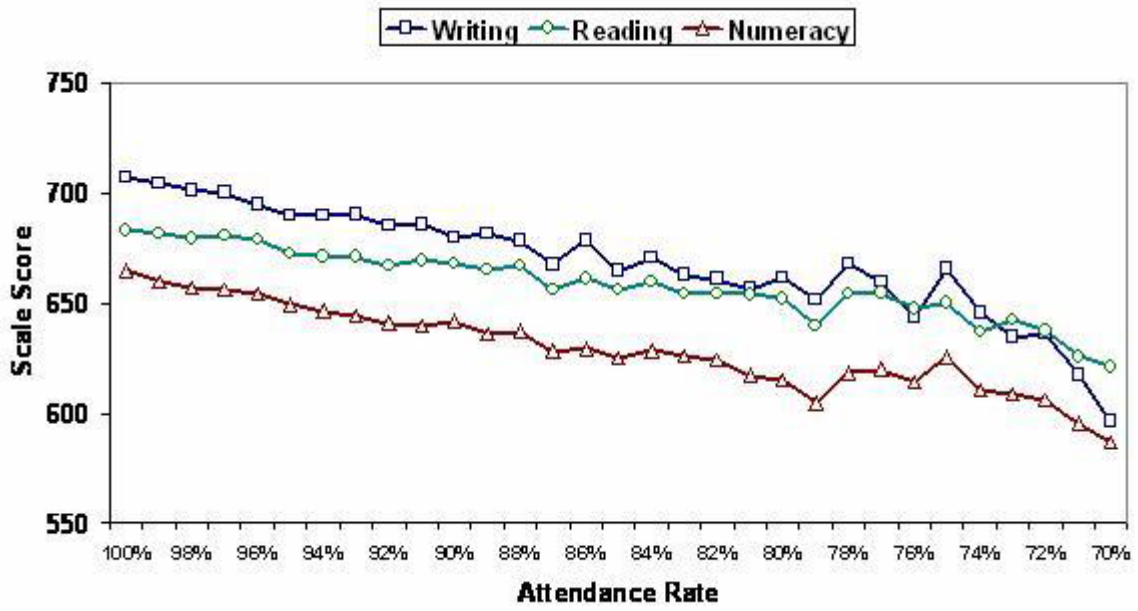
Measures of student outcome

(a) *Student attendance*

Attendance is an important indicator of school performance. Schools that are successful in having almost all their students in regular attendance are meeting a basic requirement of effectiveness. High attendance is a necessary, though clearly not a sufficient condition for success at school.

Student attendance, like student achievement, is influenced by school intake characteristics and school environment. High rates of attendance are found in schools that enrol high proportions of students from socially-advantaged backgrounds, and rates fall as the social scale is descended. This link is stronger in some jurisdictions than others, but even where the influence of SES and indigeneity is very strong, there remains a significant margin of unexplained variance. At any given level on the ICSEA scale, schools vary in attendance rates, and this suggests that some schools have more effective policies or better management than others. Schools have a strong interest in lifting rates of attendance. For these give a high prediction of success. Research by the Queensland Department of Education and Training shows a strong linear relationship between attendance and achievement, as the chart below illustrates (Queensland 2010).

Figure 2 Scale score averages for Year 7 results (2006) by attendance rate, Queensland



While student attendance rates are an important measure of school effectiveness for all schools, they may have particular relevance for the selection of secondary schools as reference schools in estimating a national resource standard. Secondary schools as a group have a wider range and lower overall rate of attendance than primary schools, as the table belows indicates.

Table 1 Government school rates of attendance

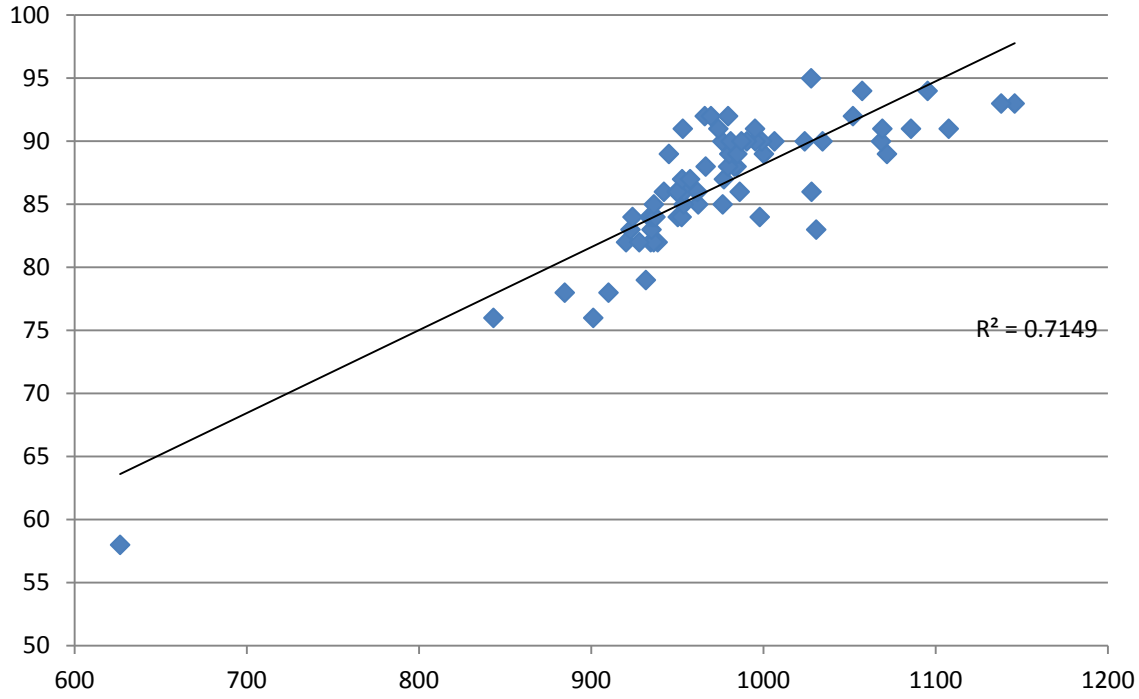
(schools not reporting rates and combined-level schools are not reported)

	Median	Mean	Standard deviation	Minimum	Maximum	Range
Primary	94.0	93.19	2.715	67	99	32
Secondary	89.0	88.69	4.087	58	100	42

Figure 3 is a snapshot of government secondary schools in one jurisdiction, showing the spread of attendance rates against ICSEA values. Almost all schools are located in the range 75% to 95%, with a mean of 86.5% and a standard deviation of 5.55%. While ICSEA is a good predictor of attendance (accounting for 71% of the variance), the spread of ICSEA values at the mean attendance rate and at about one standard deviation above the mean shows that a sample of schools drawn at either of these performance levels would cover a range of socio-economic settings.

Under the National Education Agreement (Schedule E), the Australian Government requires schools to report to their communities on student attendance, specifically (i) rates for each year-level and for the whole school, and (ii) how non-attendance is managed by the school (see, e.g., http://www.decs.sa.gov.au/quality/files/links/EssentialRequirementsAnnua.pdf).

Figure 3 Student attendance rates by ICSEA in public high schools in one jurisdiction



Aggregated rates of attendance (i.e., Years 1-10) are reported on the *MySchool* website. The base for these rates is Years 1-10, while for Victorian government schools it is Prep-Yr 10. The DEEWR website, *Student Attendance Help*, explains the concepts and data-entry procedures that schools have to follow to submit attendance data. An extract below describes concepts and data-entry procedures.

The Victorian Government reports attendance rates in School Performance Reports (see Victoria n.d.). Attendance is viewed as a measure of student engagement. Particularly in secondary school, this puts an appropriate emphasis on the measure. But at all levels of schooling, rates of attendance can be viewed as reflecting the quality of the relationship between school and community. From the angle of "community" and the angle of student "engagement", student attendance represents a suitable measure for identifying "reference schools". However, it needs to be stressed that the higher the rate that is chosen as a criterion for inclusion of a school, the narrower the intake of the school. Those schools that have high rates of attendance will tend to be high-SES and low-indigenous schools.

DEEWR, *Student Attendance Help

Attendance and enrolment data - what is it?

Schools are required to provide two sets of data to enable STATS to work out the student attendance percentage.

Enrolment data: Schools must enter in the Enrolment Data screen the total numbers of full-time students by gender and Indigenous status for each year (1-10), enrolled for the whole collection period. If any students were not enrolled for the whole period or only attended part time, they must be excluded from the data collection.

Schools will be asked to provide separately the numbers of Indigenous students, included in the above, also split male/female and for each individual year.

Attendance data: Schools must enter in STATS the attendance data for the collection period for each year (1-10), by gender and Indigenous status for all those full-time students who were enrolled for the entire collection period.

The reason why both enrolment and attendance data are needed is because the student attendance measure is expressed as a percentage. In this case the enrolment data is the denominator and the attendance data is the numerator.

*https://ssp.deewr.gov.au/ssp/help/html/sa/attendance.html

School completion

For many years, the performance of secondary schools was measured by *apparent grade retention rates* (AGRR). This measure was developed for use at a system level to monitor progress in the extent to which the population participated in extended secondary schooling. As a ratio of Year 12 enrolments to Year 7/8 enrolments (5/4 years earlier), the apparent grade retention rate has known limitations at a system level (e.g., failure to adjust for population growth, interstate or external migration, sector transfers). The apparent grade retention rate, when applied to smaller geographical units or to schools themselves—as a measure of *holding power*—also presents difficulties. Looking only at schools, the AGRR does not adjust for transfers into a school at some point between the starting year-level (Years 7/8) and the final year (e.g., Year 12). The measure cannot be applied to senior colleges or other schools that enrol students only in the senior years. Nor can it be applied meaningfully

to junior high schools (e.g., Years 7-10) or to district high schools that do not offer classes beyond Year 9 or 10.

School completion, notwithstanding measurement problems, is an important indicator of school effectiveness. Australian governments have set a target of 90% Year 12 completion or equivalent by 2015, and the school leaving age has been raised in several jurisdictions with a view to increasing young people's participation in school as well as giving them a stronger basis on which to continue in education or training.

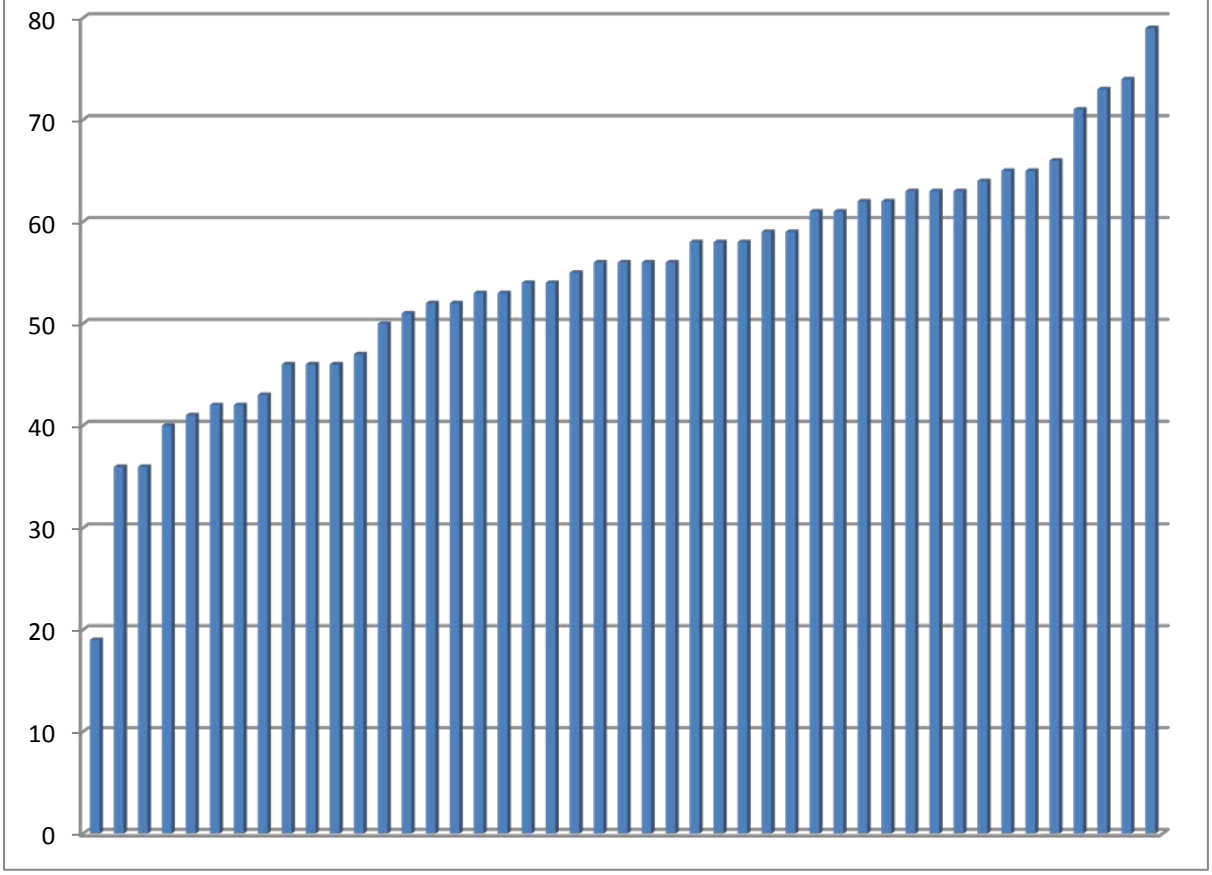
Overcoming the problem of early leaving is a major policy objective embedded in the national goals for schooling and in agreed national targets. Not all early leaving should be viewed as problematic (e.g., many early leavers start apprenticeships). However, many of the estimated 22% of young Australians who today do not complete school are low achievers or disengaged learners who are not exiting school for positive reasons and are not continuing in education or training. These young people are not a random sample of all young Australians, but are drawn disproportionately from low SES, rural and indigenous backgrounds.

High rates of school completion in schools serving low SES communities suggest high levels of student engagement, based on a range of school strategies (Lamb and Rice 2008). It would be appropriate to include "completion rates" in the set of measures for identifying reference schools for the purpose of estimating a national resource standard.

Apparent grade retention rates are not a suitable measure, for the reasons outlined above. The ABS has developed a suite of alternative measures—the School Participation Rate (SPR), the Apparent Continuation Rate (ACR), and the Apparent Progression Rate (APR) (ABS 2010). However, these are not suitable for use at a school level. An alternative approach which has been developed specifically to measure school holding power involves calculating the proportion of students enrolling in Year 7/8 and reaching Year 12. This *real retention rate* is a measure of cohort completeion rather than a grade ratio (as in the case of the AGRR). It is not affected by transfers into a school above Years 7/8, but is affected by

attrition due to transfers out as well as to early leaving. Real retention rates are reported for Victorian schools in official School Level Reports. The rates are influenced by the same factors that influence student achievement, which is correlated with "retention". Although real retention rates cannot be calculated for schools that have only upper secondary forms, most Australian secondary schools fall into the category of 7/8-12. An indication of the range of values taken by the real retention rate is given in the chart below.

Figure 4 Real retention rates, Year 7 to 12 (4-year average), Eastern Metropolitan Region, public high schools



(b) *Year 12 study scores*

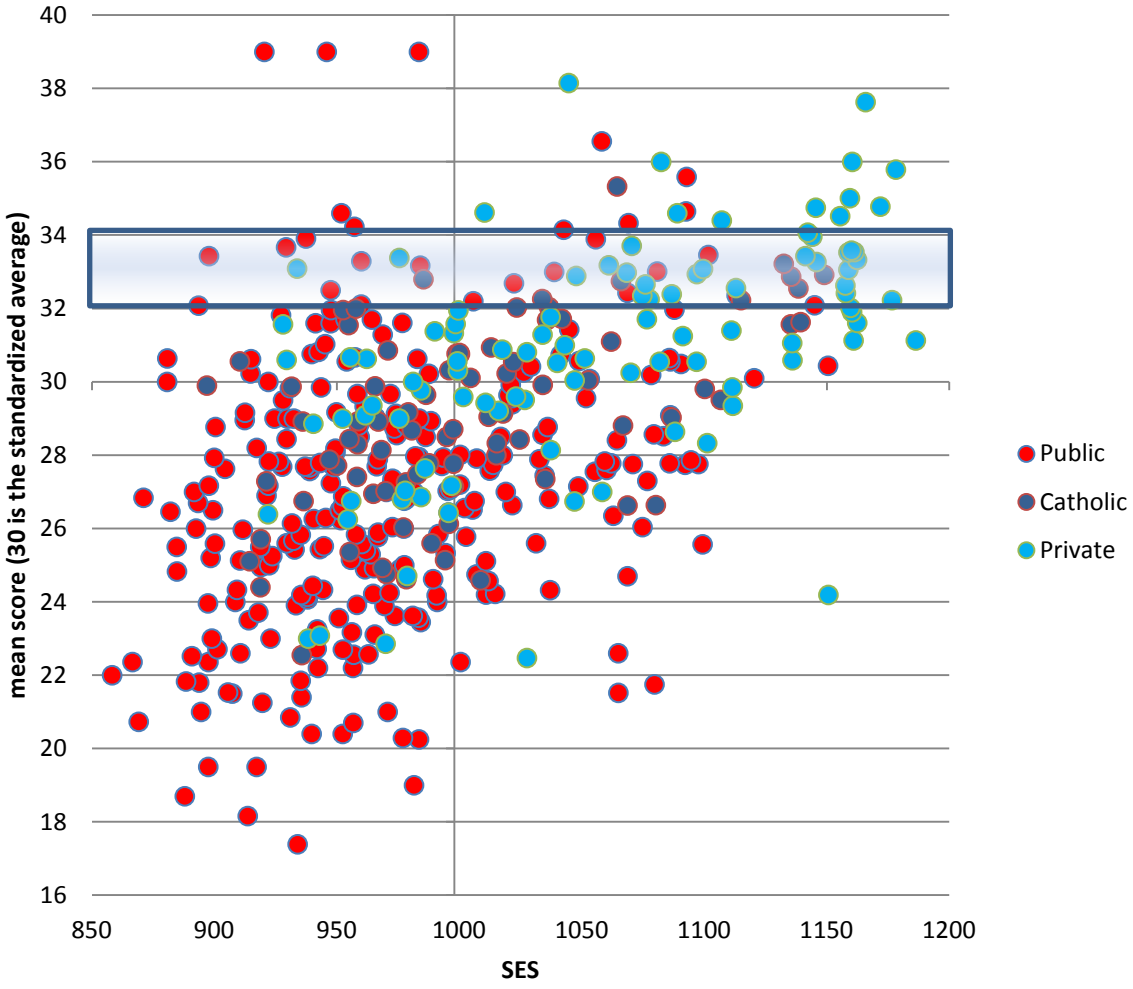
Across Australia's eight jurisdictions, there are different approaches to both curriculum and assessment in the senior years. Senior certificates are complex constructions, reflecting different student programs, most of which are tertiary-oriented, but some of which represent vocational education pathways. This is the case, for example, with the Senior Certificate in Queensland or the Victorian Certificate of Applied Learning. Assessment approaches also

differ, e.g., in the weight given to school-based assessment, how this is moderated, and the role of external exams.

While for the purposes of university selection across Australia, students are ranked on a formally common scale (ATAR), this is a construct which sits on very different approaches to certificate design, assessment and testing (e.g., the Queensland Core Skills test). Even with respect to tertiary selection, national consistency is far from complete (e.g., Queensland institutions use Overall Positions, OP).

However, even within jurisdictions, there are difficult issues which arise from seeking to use study scores as a guide to the selection of reference schools for estimating a national resource standard. Performance in senior certificate subjects that are accepted by universities for admission purposes is strongly influenced by the social intake and staff resources of schools. Depending on how high the performance bar is raised, the sample of schools becomes increasingly unrepresentative. The chart below illustrates this. In this chart, schools are ranked according to their mean study score in Mathematical Methods (university-preparatory) and the SES of students (based on home address).

Figure 5 Mean Mathematical Methods study scores by SES of school by sector, Victoria 2004

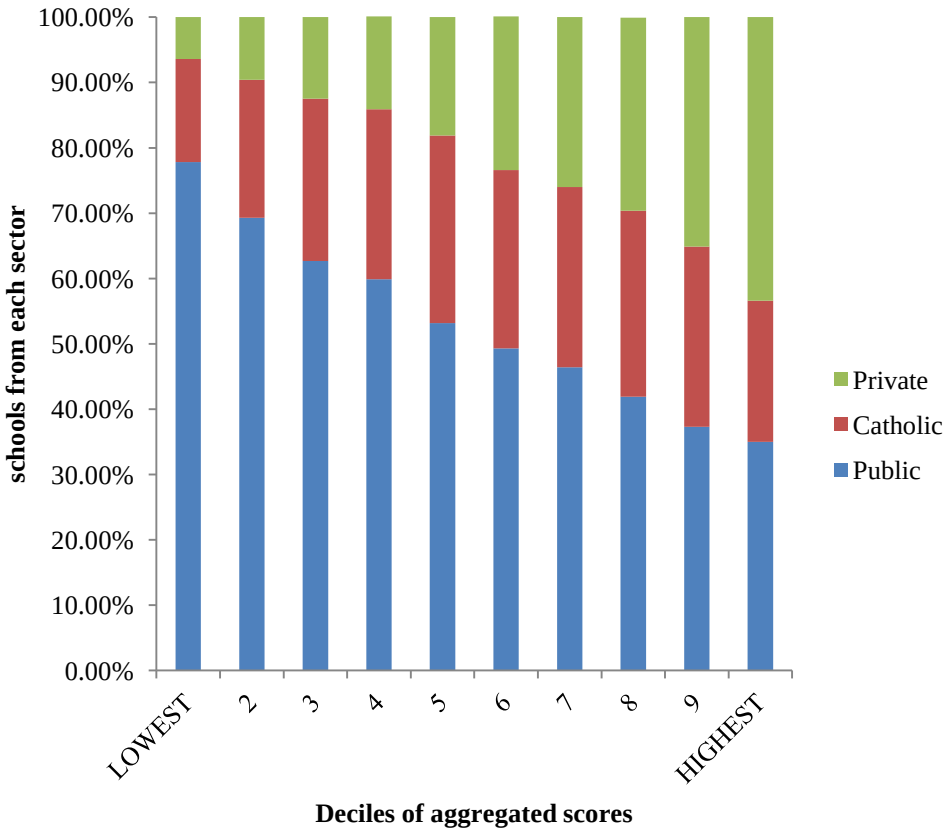


The shaded area between 32 and 34 score points represents a high performance "zone". There are 77 schools within this zone, two-thirds of which are non-government (half are private non-Catholic). Many of the public high schools that are in this zone have a relatively low SES score. Potentially they make an interesting group to study, but before doing that an adjustment would have to be made to take into account the level of enrolment in Methods. In some or perhaps all of these "exceptional" schools, the high average rates of performance could be due to low average enrolments ("trade off").

Is the difficulty of social selection overcome by aggregating across the study scores in different subjects? Performance in all studies which have significant numbers is socially biassed, so aggregating across studies does not remove the problem of bias. Indeed the bias is reinforced by the fact that more successful students take additional studies (e.g., to obtain

bonuses) and that such students more commonly come from well-educated homes. The chart below divides schools into deciles of aggregated study scores and examines the sector profile in each tenth group of performance.

Figure 6 Sector profile of schools grouped by aggregated study score Victoria, 2004



Looking first at the lowest performing schools, nearly 80% of these are public, while in the group of highest performing schools, 43% of private non-Catholic and a further 22% are in the Catholic sector. Altogether around two-thirds of schools at this level of performance are non-government.

Aggregate study scores are affected by bias from two other sources—differences in rates of completion and streaming into vocational tracks. Higher scores are gained when weaker

students leave or are channelled into vocational streams. Attrition and track allocation can artificially inflate performance levels. Schools, by contrast, that operate a policy of inclusiveness and seek to maximize participation in academic tracks run the risk of lower overall performance. Basically the problem with this measure is that it is not "whole of cohort".

(c) Australian Tertiary Admission Rank (ATAR)

The Allen Consulting Group report includes in a list of possible student outcome measures the average rank of Year 12 students as calculated for tertiary admissions. ATAR is a normative or ranking measure. It relates to all school leavers who apply for a tertiary place through a tertiary admissions centre.

ATAR can be regarded as a measure of student achievement, even though its purpose is to rank qualified applicants for selection. Institutions themselves view ATAR as a measure of prediction of success in tertiary studies rather than a general measure of achievement. ATAR scores (ranks) are influenced by the choice of subjects taken by students as these are weighted to reflect relative difficulty. The weightings transform the original study scores so that a score is either adjusted upwards (in some cases by a substantial margin) and in other cases adjusted downwards. Typically subjects, such as university-preparatory maths, physics, and some traditional humanities, are scaled up, while some of the more modern humanities and business studies are scaled down. The average ATAR for a school is thus a reflection of (a) the pattern of subject choice, and (b) the relative academic success of students. Schools that have high ATARs have more successful students, but also students who taken certain subjects more often than others.

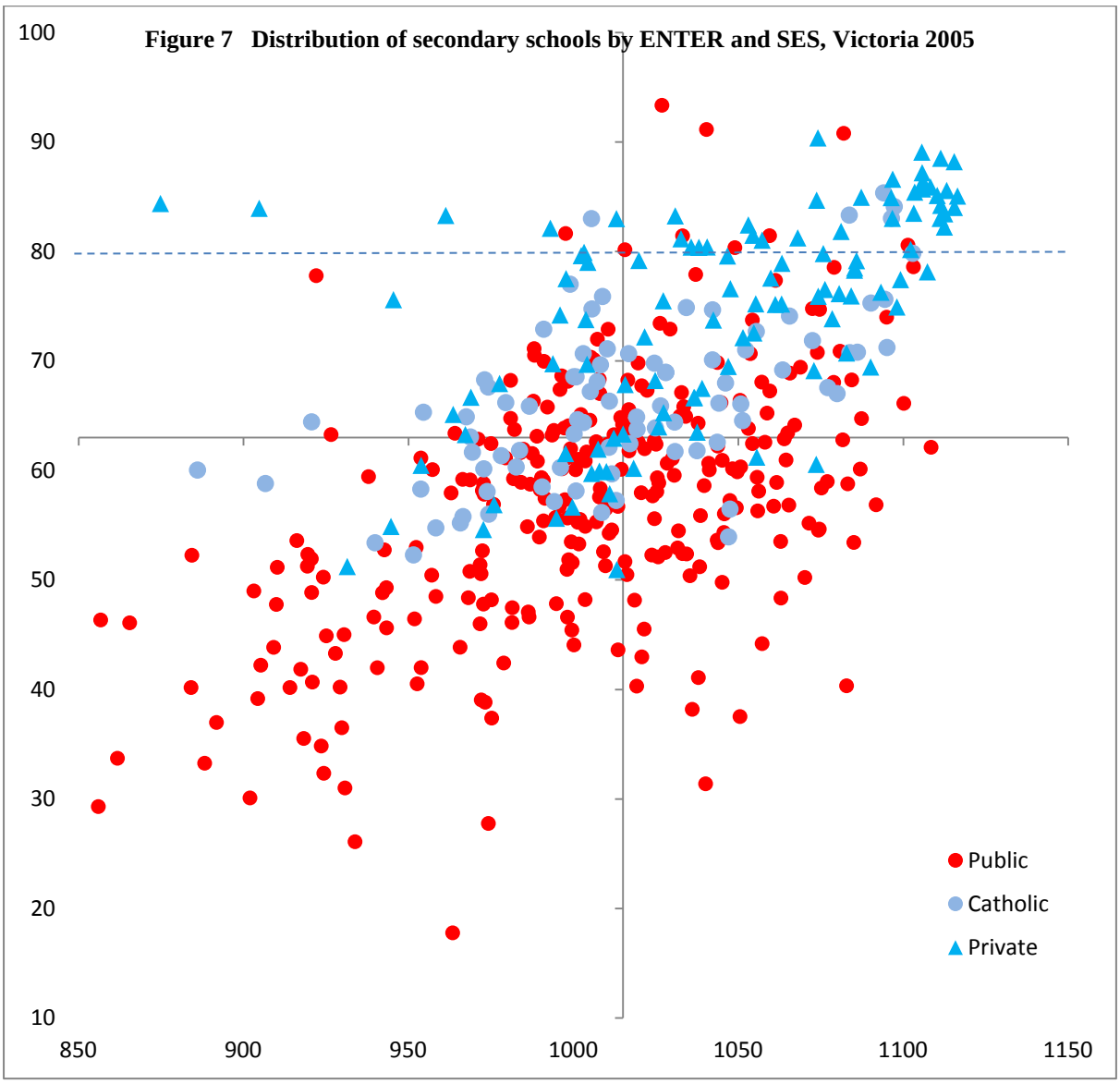
A high ATAR does not necessarily mean that a school is high performing. A school may achieve a high rank because relatively few students complete school or because many does so in a vocational strand which does not generate a tertiary selection score. Many Year 12 students do not seek a tertiary place and do not make an application to a tertiary admissions centre. A high ATAR would not reflect their presence in Year 12, and if there are many such students, the rank would be a misleading guide to school performance.

Basically a *whole of cohort* measure of school performance is desirable rather than one which captures only a sub-set of the cohort, even if this is a large sub-set. For example, a measure of school completion reflects the extent to which a school succeeds in engaging all of its students, even if there is a wide range of achievement in the completing cohort. Similarly student attendance relates to the whole cohort rather than a group.

Schools with high ATARs are not an unbiassed sample of all secondary schools in terms of social intakes and resource levels. They are schools which demonstrate a high degree of competitive advantage, based on selectivity (academic and social), resources and expertise (specialist staff), and academic focus (specialization in the curriculum).

A snapshot of Victorian schools ranked by ENTER and by SES is provided in the chart below. There is a close association between these two series (0.6), even without any adjustment for school size or location. High-performing schools are displayed in the first quadrant (upper right). These are in the main non-government schools.

Figure 7 Distribution of secondary schools by ENTER and SES, Victoria 2005



Taking a rank of 80 as a threshold would produce a selection of mainly private non-Catholic schools (74% of the total of high performers), a small group of Catholic secondary schools (9%) and a mix of selective-entry and local public high schools (17% of the total).

(d) VET unit completion

The provision of nationally accredited modules of VET in the curriculum is one of the most important innovations in upper secondary education in Australia since the mid-1990s. The aims have been to improve student motivation and engagement, to develop vocational skills and employability, to increase school completion, and to ensure better employment, study or training transition.

The number of students completing a unit of competence is a measure of the extent to which schools have succeeded in diversifying their upper secondary curriculum to provide a broader range of choice—particularly for academically less able students—and the success of their students in undertaking accredited VET. In addition, the measure balances the perspective furnished by ATAR which represents school success in terms of preparation for and relative access to tertiary studies.

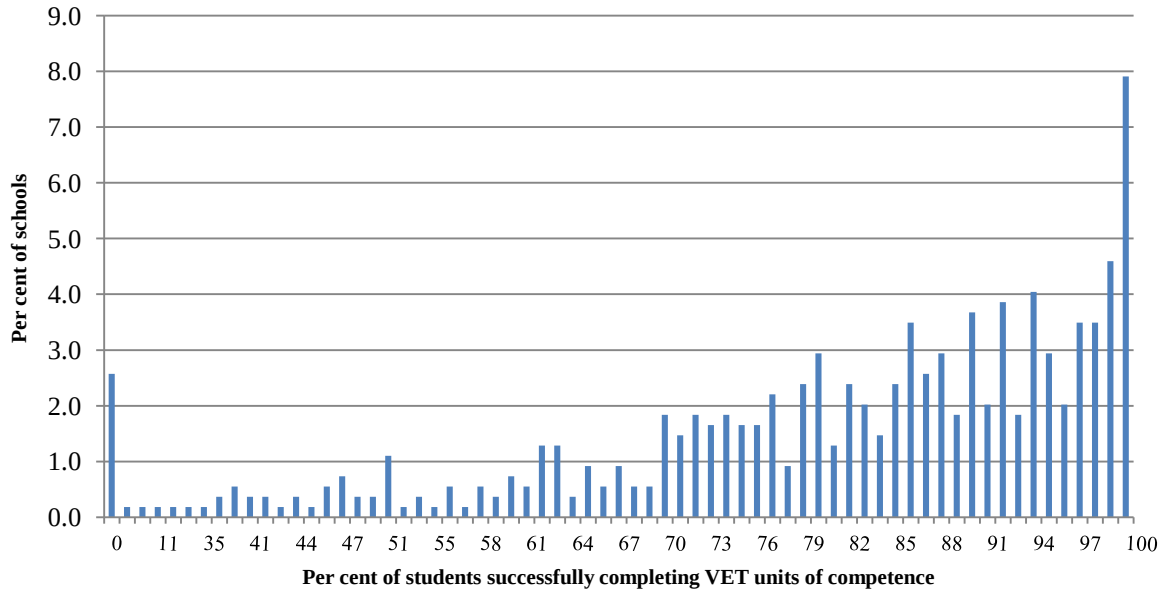
However, there are several issues with the measure as included in the Allen Consulting Group report that need to be addressed. Firstly, the measure—"units completed"— refers to the *number of students* achieving a unit of competence. This will vary according to a range of factors. These include the size of the student cohort in a school, the relative importance of VET certificates as a component of total offerings, the quality of VET provision, and the quality of student effort.

A measure of achievement of VET competence should be adjusted to reflect the relative extent of participation in VET (not simply the number of students achieving units of competence). As currently expressed, the measure would give equal rank to a school in which the majority of students who attempted VET units successfully completed them and a second school in which only a minority successfully completed their units. This is because the measure is not adjusted for participation in VET.

The *MySchool* website does not report a VET "participation" measure. It only reports "achieved units of competence". In at least one jurisdiction, the curriculum authority does report the number of students *enrolling in a VET certificate* (VCAA 2010) and also the *percentage of students who complete units of competence*. This second measure (*percentage completion*) is the more relevant, but there are definitional issues to be addressed. It is unclear whether this measure is a ratio of units enrolled to units completed, or whether it refers to the percentage of *students* who successfully completed one or several units of competence.

The distribution of schools on the measure of *percentage completion* is skewed upwards, with over 60% of schools in Victoria recording at least 80% completion rates; 39% have rates of 90-100% completion (as the chart below shows).

Figure 8 Victorian secondary schools and VET completions (%)



As students may attempt multiple units of competence, this measure may reflect the success of the same student on several modules as well as the success of many students taking only one module.

Further research is needed to determine how the percentage completion measure is calculated (is the unit an enrolment or a student?) and whether this measure could be produced nationally.

Given the objectives of VETiS, it may be preferable in light of the issues discussed above to capture the success of schools in delivering VET by using measures of the *impact of VET* rather than direct measures of enrolment or module or award completion. One intended impact of VET is *school completion*. Another is *post-school transition*, especially to further training and employment.

(e) Post-school transition

Several States collect data on student destinations on leaving school. Victoria and Queensland run comprehensive student surveys, involving direct contact with school leavers. Western Australia collects data on both student intentions and student destinations. The methodology of *On Track* (the Victorian program) has been applied in a number of large cohort studies in New South Wales and in pilot studies in South Australia.

On Track was developed to answer the need expressed by school principals for a broad reflection of student outcomes as compared to newspaper league tables which reported exam results. A well-constructed survey captures the full range of education, training and employment outcomes, and gathers additional data, such as occupations, hours of work, and course attended (for those exit students who are in further study). In the past, *On Track* also gauged student satisfaction from the angle of the school leaver (as compared to school-based surveys of current students). *On Track* and *Next Step* programs write reports for all schools as a resource for curriculum planning, professional development, and the organization of counselling and other support services.

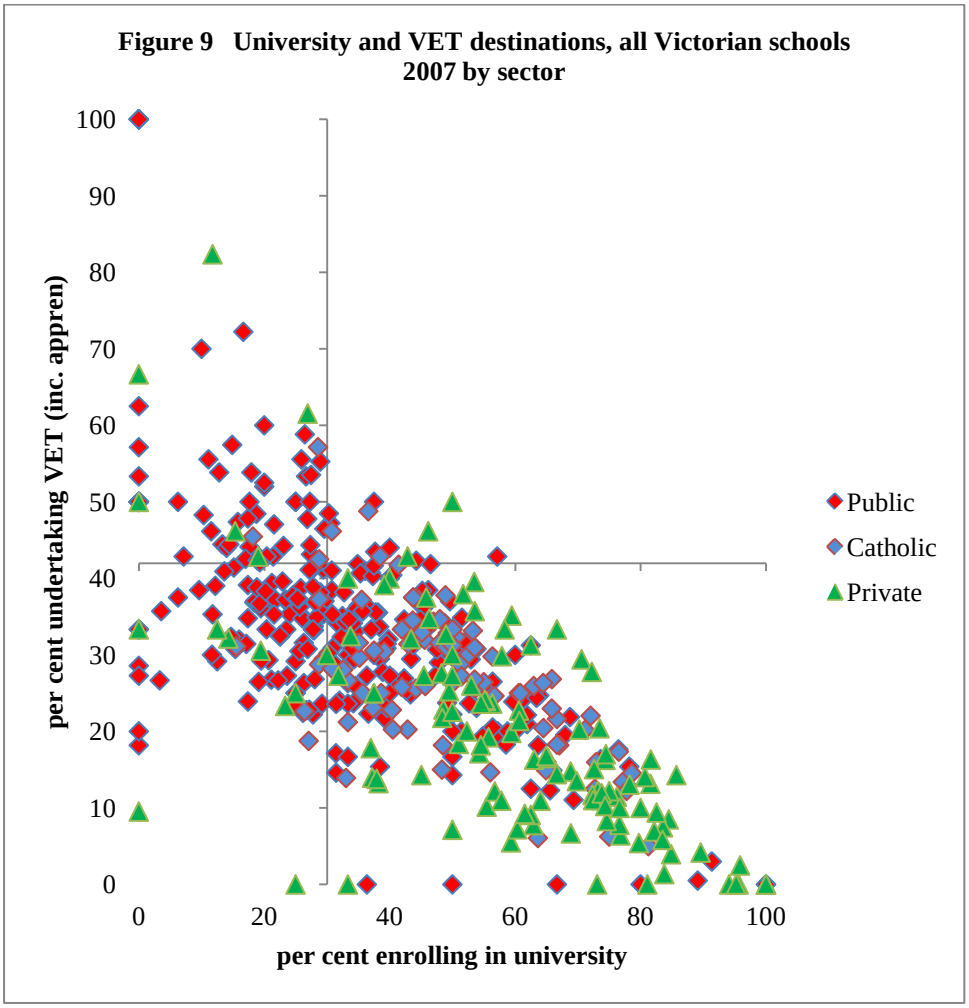
In theory, measures of post-school destinations are "whole of cohort", but in practice this may not be so. Response rates vary, and are problematic for the hard-to-reach groups who leave school early. Also participation is voluntary, and sizeable numbers of students do not agree to participate. Thus destinations surveys do not reach all school leavers, and the school leavers that they do reach are not an unbiassed sample of all leavers. School-to-school variability in response rates and participation rates weakens the potential value of destinations measures for expressing the quality of performance of individual schools.

However, this does not mean that results from data on destinations from Western Australia, Queensland and Victoria could not be used in the selection of reference schools. One question concerns the design of the measure of post-school destinations, a second question concerns response rates/participation rates, and a third question concerns whether the measure as designed would capture a good range of schools.

It matters how the destinations measure is constructed. For example, if a "good" school is defined as one that reaches a high threshold of students entering university, this will capture a large group of private non-Catholic schools, selective-entry state high schools, Catholic independent schools, and public high schools located in upmarket suburbs of Australia's major cities. By contrast, if a measure of destinations is used which focusses on VET—middle-level courses, basic, apprenticeships, traineeships—this will net a large group of public high schools (generally serving middle to low SES communities).

In the chart below, Victorian secondary schools are ranked on each of these constructs—the proportion of students who enrolled in university (horizontal axis) and the proportion who undertook VET, either employment-based on campus-based.

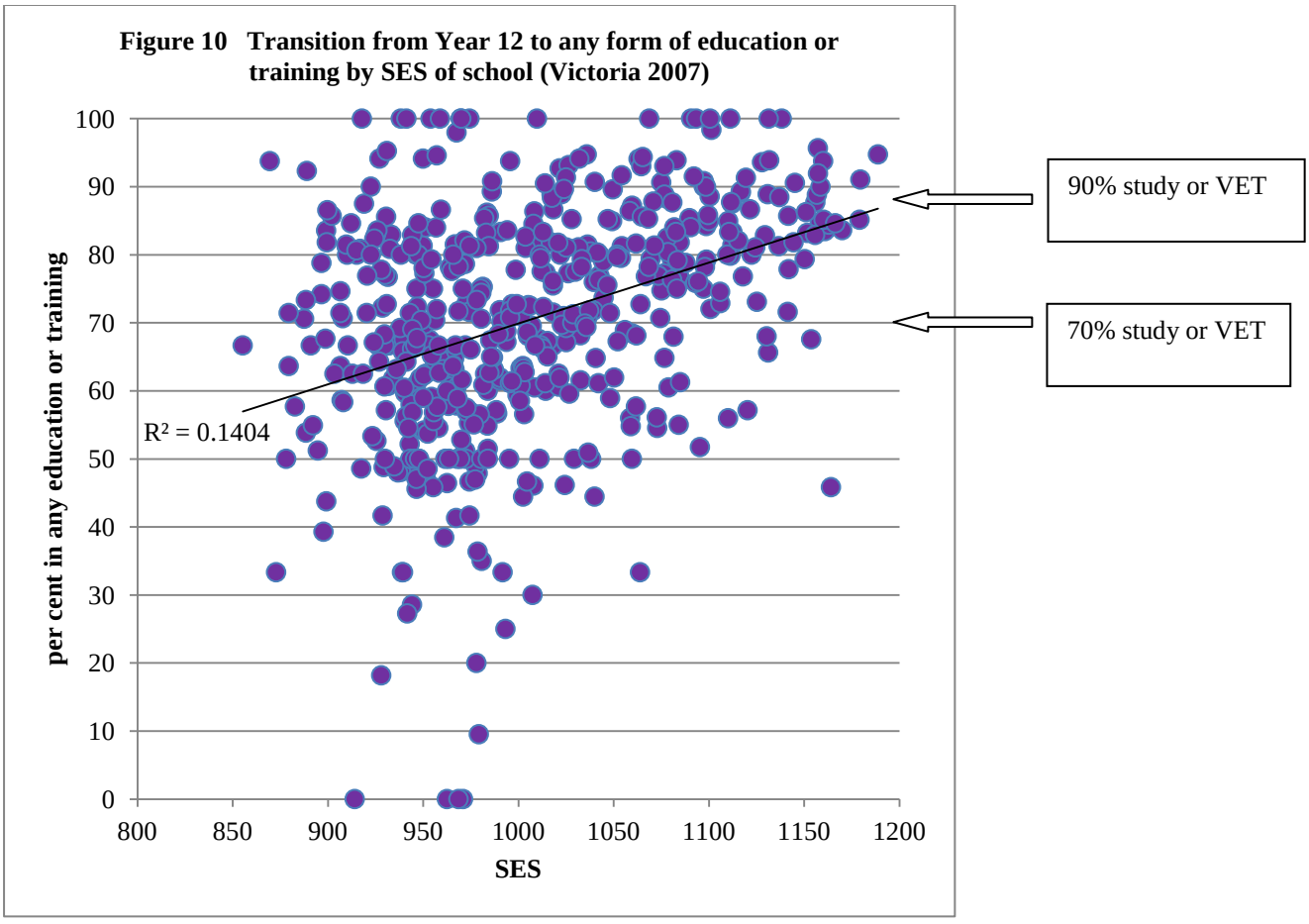
Figure 9 University and VET destinations, all Victorian schools 2007 by sector



As will be apparent, there is a large, negative correlation between the two series (*r=-0.7112*). The higher the proportion of students enrolling in university, the lower the proportion undertaking VET. The point of this analysis is to show that a "destinations" measure needs to be comprehensive if it is not to favour one group of schools over another, and thus import both a social and an academic bias into the reference schools.

One approach is to simply aggregate all destinations that involve education or training—university, TAFE diploma/advanced diploma, Certificates 1-4, apprenticeship and traineeship. Such a composite measure does not discriminate between sector of activity, and as a result the measure captures a large group of schools from across different sectors and different geographical contexts. This spread in schools is reflected in the chart below, which relates the proportion of school completers undertaking any form of education and training to the SES of the school (based on SEIFA of home addresses).

Figure 10 Transition from Year 12 to any form of education or training by SES of school (Victoria 2007)



Although there is a positive association between "study or training" and school SES (*r=0.3746*), it is important to look at the spread of schools at different levels of performance on the "study or training" scale. At the 70%, 80% and 90% level of study or VET transition, there is a wide social range of schools. The sector mix of schools will vary at these different levels.

As the choice of schools is intended to access the different costs of high quality schooling, it would be important to draw from across the social spectrum. For the further down the social scale, the greater the intensity of VET transition—and therefore the more relevant the costs associated with VET. The further up the social scale, the greater the intensity of university transition and therefore the more relevant are costs, such as small specialist classes or the employment of experienced, specialist and expensive staff.

Conclusion

While the measures listed in the ACG report are broad in range, not all are equally suitable or robust. Before turning to the most suitable measures, it is important to stress key points discussed earlier in this report.

First, the thresholds applied to different measures (e.g., 80%+ NAPLAN) should not be set purely with regard to *high* performance. Focussing on high performance runs the risk of selecting schools whose high performance on a particular measure may not be related to the level of resources used in the school, but reflects advantages of home background or selectivity of intake (or both).

The aim should be to set thresholds which select schools experiencing different cost pressures, all of which have to be addressed by a national resource standard. From within this range, some schools will reach the proposed NAPLAN threshold, for example, but others will not. The validation study would aim to estimate the level of resources required to enable all schools to reach this standard.

Australian schools perform well by international standards, and there is no reason for choosing from *amongst the many good schools* only those that perform *at high thresholds*. Moreover, as we have stressed, narrowing the focus to include only these schools is likely to result in the inclusion of some that are unsuitable and the exclusion of some that should be included.

Secondly, performance information can be used to help estimate a resource standard in a different way than choosing a sample of high quality schools. A purposive sample of schools in different characteristic settings has the advantage of reflecting multiple cost factors, while giving access to schools that (a) exceed expectations, (b) operate at benchmarks, and (c) under-perform expectations.

Such an approach enables an assessment to be made of what *it costs to make a difference* (Lamb 2011). What cost differences are associated with variation in performance of schools serving similar communities? Are these differences connected, for example, with higher staffing costs, different programs, services and activities, differences in levels of teacher release time for professional development, or targetted small class sizes, team teaching, etc.?

Thirdly, the choice of reference schools must have stakeholder credibility. If schools are chosen using performance criteria (measures and thresholds) that net a sample which is not representative of all schools, doubt will be cast on the process and the result of the estimation process.

Current reviews of school funding in Victoria, Western Australia and Tasmania (and a review about to be launched in New South Wales) make the national review doubly sensitive. The principals of schools serving poorer communities feel that their schools are disadvantaged by current models and need reassurance that the particular set of costs associated with their challenges will be reflected in a national resource standard. The *costs of improvement*—what these schools need to get significantly better outcomes—is especially important. This is not to say that the principals' own assessments should be uncritically accepted.

Recommendations

It is recommended that a purposive sample of schools be selected which is (a) reflective of *characteristic settings* across Australia, and (b) reflective of a *range of performance* on each of the following measures.

Schooling level	*Measure*	*remarks*
Primary	NAPLAN	
	Student attendance	
Secondary	NAPLAN (Yr 9)	
	Student attendance	
	School completion	Based on real retention rates
	Post-school transition	Exit students in any education or training
Combined	Mix of measures	Varied according to enrolment profile
(e.g., District Highs)		

A number of measures recommended for inclusion are not available nationally—real retention rates, post-school destinations—or require a modification to current reporting practice (student attendance in Victoria). The reporting of student attendance rates in Victoria can be adjusted to align with practice across other jurisdictions. Real retention rates can be calculated by schools, using admissions records. Post-school destinations should be employed in selecting schools in Western Australia, Queensland and Victoria as this will give access to cost factors associated with the delivery of a broad range of post-compulsory programs and differences in curriculum emphasis. However, schools should only be included that have high participation and response rates, and the measure of destinations should be reflective of *all* school leavers, not only those who complete Year 12.

For secondary schools, NAPLAN (Year 9) provides the only set of achievement measures that can be consistently applied at a national level. Year 7 results are relevant, but as a means of adjusting for the achievement levels of children at the time of their entry to secondary school (see for intake adjustment in publicly reported measures of school performance in Victoria http://www.eduweb.vic.gov.au/edulibrary/public/commrel/policy/Blueprint2008/iamexplanation.pdf). In the case of Western Australia, it is possible to relate senior high school performance in Year 9 with the achievement of children in feeder primary schools in Year 7. Adjusting measures of Year 9 performance for achievement levels at intake captures the value added by schools, and this is a valuable measure of the quality of a school's work. This represents the difference that schools make, and should be preferred to relying on the absolute level of performance.

Measures of literacy and numeracy do not provide a very broad picture of student achievement across the curriculum, whether in primary or secondary school. Assessment information relating to Essential Learnings—e.g., in South Australia, Queensland, Tasmania and Victoria—could be drawn on to provide a fuller picture, though based on teacher assessments.

For sources, see:

http://www.qsa.qld.edu.au/downloads/approach/qsa_assessment_overview.pdf;
http://www.sacsa.sa.edu.au/index_fsrc.asp?t=EL;
http://www.education.tas.gov.au/school/curriculum/assessment/tcf;
http://www.education.tas.gov.au/curriculum/standards.

The different curriculum frameworks designed for years of schooling up to Year 10 vary across jurisdictions, and reporting of school performance also varies. As the new national curriculum is implemented, there will be a progressive alignment of these standards, and a nationally consistent set of measures which are curriculum-based will become available (for a discussion of the transition, see Howes (nd).

Whether progress towards implementing the national curriculum at upper secondary levels will lead to a consistent approach to assessment, including the role of external examinations, the relative weight given to school-based assessments, and the participation of all students in assessment, is difficult to say at this stage in the implementation process.

The ACG report mentions parental satisfaction as a possible measure. Some systems (e.g., Victoria, Queensland, South Australia) do have a survey program in place, but response rates are not reported. For a sample of reporting from a parent opinion survey, see the exhibit below.

Parent Opinion

School no: 7595
School name: Bendigo Senior Secondary College
Explanation: A self-completion questionnaire to obtain parents' opinions of the school. Parents rated a number of statements using a scale where 1=strongly disagree & 7=strongly agree. The statements were grouped into the following categories.
Source of data: Data scanned from questionnaires completed by parent

Variable No.	Name	Year	School mean	State Secondary median
1	School Improvement	2008	5.44	5.07
		2009	5.28	5.10
		2010	5.52	5.12
2	Approachability	2008	5.52	5.12
		2009	5.40	5.09
		2010	5.43	5.11
3	Teacher Morale	2008	5.32	4.89
		2009	5.23	4.90
		2010	5.48	4.94
4	Parent Input	2008	5.09	4.84
		2009	5.16	4.87
		2010	5.23	4.84
5	Stimulating Learning	2008	5.26	4.82
		2009	5.03	4.84
		2010	5.32	4.88
6	Behaviour Management	2008	5.16	4.71
		2009	5.17	4.74
		2010	5.32	4.74
7	Reporting	2008	5.52	5.17
		2009	5.32	5.15
		2010	5.34	5.13
8	Learning Focus	2008	5.91	5.22
		2009	5.69	5.22
		2010	5.93	5.24
9	Transitions	2008	5.43	5.24
		2009	5.32	5.25
		2010	5.57	5.28
10	Extra Curricula	2008	4.99	4.65
		2009	4.79	4.65
		2010	4.86	4.64
11	Homework	2008	5.02	4.57
		2009	4.73	4.56
		2010	4.93	4.61
12	General Satisfaction	2008	5.92	5.32
		2009	5.54	5.29
		2010	5.91	5.32
13	Student Safety^	2008	6.26	5.44
		2009	6.48	5.40
		2010	6.23	5.45
14	Classroom Behaviour^	2008	4.46	3.48
		2009	4.51	3.49
		2010	4.69	3.52
15	Connectedness to Peers	2008	5.89	5.67
		2009	5.87	5.67
		2010	6.08	5.68
16	Student Motivation	2008	4.76	4.78
		2009	4.73	4.77
		2010	5.23	4.83
17	Social Skills	2008	5.63	5.40
		2009	5.59	5.42
		2010	5.91	5.43
18	School Connectedness	2008	5.05	4.94
		2009	4.91	4.95
		2010	5.40	4.98

There is potential is using measures of parental satisfaction, but much better information is needed regarding which parents participate, how many, and how consistent the ratings are over several years.

Student opinion is also a potentially valuable source of information about quality of school performance. There are known patterns in student responses to a range of questions on quality of instruction, school climate, and classroom climate, and high levels of student

participation are potentially available to such surveys. However, there is currently no nationally consistent approach to surveying student opinion, and this presents difficulties in comparing schools from different jurisdictions.

References

ABS 2010. *Schools 2009*. Appendix 3 Alternative Measures of Engagement in Secondary Education (Catalogue No. 4221.0)

Howes, D. (n.d.) *Transition to the Australian Curriculum in Victoria*

http://www.vcaa.vic.edu.au/vcaa/natcurric/Transition_to_the_Australian_Curriculum_in_Victoria.pdf.

Lamb, S. (2011) Private communication

Lamb, S. and Rice, S. 2008. *Effective Strategies to Increase School Completion Report* (Melbourne: Department of Education and Early Childhood Development) (www.education.vic.gov.au/sensecyouth/careertrans/)

Queensland 2010. *Every Day Counts.*

(http://education.qld.gov.au/everydaycounts/research.html)

VCAA 2010. Post *Compulsory Completion and Achievement Information, 2010*

http://www.vcaa.vic.edu.au/vce/statistics/schoolstats/postcompcompletiondata-2010schools.pdf

Victoria (n.d.). *Transition and Accountability Principles* (PowerPoint presentation)

(http://www.education.vic.gov.au/management/schoolimprovement/performancedata/performancereports.htm)