

Using value-added measures to identify school contributions to student learning

Centre for Education Statistics and Evaluation

The Centre for Education Statistics and Evaluation (CESE) has developed a set of value-added (VA) measures for NSW government schools.

VA measures are typically used by schooling systems to indicate the contribution that a school makes to student learning, over and above the contribution made by the average school.

VA measures examine student progress over a specific time period, and adjust for factors that are outside the control of schools (such as students' socio-economic status). This provides a fair and accurate indication of the effectiveness of schools.

The development of these VA measures is based on the latest international research and experience. The VA measures have been reviewed by leading experts in education.

Key features of school VA measures:

1. All schools add value. VA measures indicate a school's contribution to its students' learning, relative to the contribution of the average school (after adjusting for differences in student and school characteristics).
2. These measures are not about student achievement levels, but about shedding light on the effectiveness of schools.
3. The VA measures do not tell us *why* some schools add more value than others. However, they are a good starting point for more in-depth understanding of 'what works'.



CESE has developed a suite of value-added (VA) measures of school effectiveness to overcome limitations of other commonly-used indicators of school performance¹. VA measures indicate the contribution that a school makes to the learning of its students, over and above the contribution made by the average school.

The VA measures are based on the latest research and experience of education systems world-wide that are increasingly implementing similar measures². The development process was reviewed and endorsed by two educational experts: Dr. John Ainley (Australian Council for Educational Research) and Professor Magdalena Mok (Hong Kong Institute of Education).

Why do we need value-added measures?

Measures commonly used to estimate the impact that schools have on student learning include:

- Absolute performance scores at a particular level of schooling, for example the average score on the Year 3 NAPLAN Reading test, as published on the My School website.
- Gain score measures (or growth measures), for example the difference in students' achievement between NAPLAN Year 3 Reading and NAPLAN Year 5 Reading tests. The My School website as well as the Department's analytical tool – SMART – both use gain scores to compare student progress relative to students with similar starting points or similar socio-economic backgrounds.

Performance scores, and to a lesser degree gain scores, do not take account of factors that we know impact on student results. Schools with higher socio-economic status (SES) tend to have both higher absolute performance at a point in time and higher growth between two points (as shown in Figures 1 and 2). Because of these relationships, evaluating a school's contribution to learning based solely on absolute or gain measures puts lower SES schools at a disadvantage.

The new VA measures take into account those contextual factors (both school- and student-related) that impact on students' learning and that are largely beyond the control of

schools. They help to identify schools that make a larger than average contribution to students' learning. The VA measures can be used as the basis for further understanding of 'what works' to provide sound evidence for educators and policy makers to continually improve teaching and learning.

What value-added measures are being produced?

School effectiveness is complex, and schools may add more or less value to student learning at different points of schooling. For example, a school might be particularly effective at producing learning gains for students in the early years of schooling, but be less effective in later years. To enable schools to identify their individual strengths and weaknesses, separate VA scores are produced. These scores cannot be combined to produce one measure for each school. The scores are:

- **VA3-5**, measuring the contribution between Year 3 and Year 5. Average Reading and Numeracy NAPLAN scores are used for both years.
- **VA5-7**, measuring the contribution between Year 5 and Year 7. Most students change schools in Year 7 to attend a designated secondary school. This measure tracks students to their new school and attributes all growth between Year 5 and Year 7 to the primary school, because students will have only attended secondary school for a small number of months before sitting the NAPLAN test. Average Reading and Numeracy NAPLAN scores are used for both years.
- **VA7-9**, measuring the contribution between Year 7 and Year 9. Average Reading and Numeracy NAPLAN scores are used for both years.
- **VA9-12**, measuring the contribution between Year 9 and Year 12. Three different VA9-12 scores will be produced: the first examines the learning progress in English from Year 9 (as measured by scores on the NAPLAN Year 9 Reading test) to Year 12 (as measured by scores on HSC English subject tests); the second examines the progress in numeracy (from NAPLAN Numeracy test in Year 9 to HSC maths subject tests in Year 12); and the third examines the overall progress from Year 9 (as measured

Figures 1-2 – Absolute and gain score measures are both related to school SES



Source: Data reported on SMART for NSW government primary schools (2012).

Figure 3 – A typical presentation of value-added scores for three schools



by average scores on NAPLAN Reading and Numeracy tests in Year 9) to Year 12 (as measured by Tertiary Entrance Scores³ in Year 12).

A VA measure indicating contributions to learning from Kindergarten to Year 3 is also being considered, in order to produce a more complete picture of the contributions made by primary schools.

Scores will only be calculated using students who attended the same school in both years, with two exceptions. The first is VA5-7, where students changing to a secondary school in Year 7 will be aligned to their originating primary school. The second is VA3-5, where students changing schools to attend an Opportunity Class in Year 5 will be aligned to their originating primary school.

Interpreting value-added scores

All schools add value to student outcomes. A VA measure is an estimate of the contribution a school makes to student learning, over and above the average school, after adjusting for background factors. Almost all government schools will be assigned several VA scores⁴. A positive VA score indicates that a school is contributing to the learning of its students by more than the average school. A negative VA score indicates that a school is contributing to learning by less than the average school. This difference in effectiveness might be due to differences in teacher practices in the school, differences in school leadership or organisation, the way in which the school engages the community, or other (unmeasured) factors.

While VA scores provide a better measure of school performance than absolute or gain score measures, they are not immune to measurement error. To acknowledge this, VA scores will always be presented with confidence intervals. A confidence interval is an indication of how sure we are about the estimated score. Figure 3 shows an example of how a VA measure might be presented. The pink dot represents our best estimate of the effect of each school, and the line either side of the dot represents the confidence interval. There is a 95 per cent chance that the true effect of the school is somewhere along the line.

If all of the confidence interval line is above the zero point in the graph (such as School A in Figure 3), then a school has a *statistically significantly positive* VA score. We are quite sure that the school is adding more value than the average school. If all of the confidence interval line is below the zero point (such as School B in Figure 3), then a school has a *statistically significantly negative* VA score. This does not mean that the school is not adding value to its students – just that the school is adding less than that added by the average school. If the confidence interval line overlaps the zero point (as in School C in Figure 3), then we cannot be sure whether the school is adding more or less value than the average school. For these schools, we say that their value-added scores are *not significantly different* to the average school. Again, this does not mean that the school is not contributing to the learning of its students.

VA scores are measured in standard deviations of the scores for the later year measured. For example, a VA3-5 score of 0.2 indicates that the school contributes an additional one-fifth of one standard deviation to the achievement of its students in average Year 5 Reading and Numeracy scores, compared to the average school's contribution. Table 1 shows the equivalents in NAPLAN points and bands for selected VA scores. If a school has a VA3-5 score of 0.3, this means that the school contributes an additional 22 points (translating to about two-fifths of a NAPLAN band) to its students' Year 5 scores, compared to the scores of students in an average school.

Table 1 – VA scores in NAPLAN equivalents

VA score	NAPLAN points	NAPLAN bands
0.1	7	0.1
0.2	15	0.3
0.3	22	0.4
0.5	37	0.7

Note: NAPLAN points displayed in this table are an average of Reading and Numeracy scores, calculated using performance of matched students only in NSW government schools. These estimates may differ slightly based on which scholastic and calendar years are examined.

Features of the NSW Public Schools VA measures

To ensure that valid and reliable interpretations can be drawn from the VA scores, the following features are incorporated in the models:

Adjusting for external contextual factors

To produce fair measures of school effectiveness, VA scores adjust for factors that are outside the control of schools. In order to be included in the model, a contextual factor must have a substantial, consistent effect on student outcomes, and must be largely out of the control of schools. The contextual factors that meet these criteria are:

- **Students' socio-economic status** – based on parental education and occupation as provided on enrolment records.
- **Students' Aboriginal or Torres Strait Islander status** – based on parent or student identification on enrolment records.
- **School socio-economic status** – based on the Department's Family Occupation and Education Index (FOEI), which is used for internal resource allocation⁵.
- **Opportunity Class (OC) enrolment.** [VA5-7 only]
- **Students' gender.** [VA9-12 only]
- **Fully academically selective schools.** [VA7-9 and 9-12 only]
- **Co-educational, girls only, or boys only schools.** [VA7-9 and 9-12 only]

As well as these external, contextual factors, all VA measures also adjust for student prior ability (based on the average scores in Reading and Numeracy in the previous NAPLAN tests). This is because the focus of the value-added measures is on the progress of students between two points.

Pooling data across time to improve reliability

A value-added estimate for any one year contains a lot of uncertainty. To make the estimates more reliable, data is pooled across two measurement periods. For example, a VA3-5 score

reported for 2014 will be based on two cohorts of students – one moving from Year 3 in 2011 to Year 5 in 2013, and the next cohort moving from Year 3 in 2012 to Year 5 in 2014. This ensures that estimates are reflecting persistent differences in school performance, rather than normal variation.

Multilevel modelling to help produce more reliable and accurate school effect estimates

This statistical approach recognises that, as well as students differing from each other within a school, the types of students that one school enrolls might be very different from the types of students another school enrolls. This provides more reliable and accurate school effect estimates. It also allows us to better distinguish between the effects of students' characteristics (for example, whether a student is Aboriginal or Torres Strait Islander) and the effects of school characteristics (for example, whether a school is single-sex).

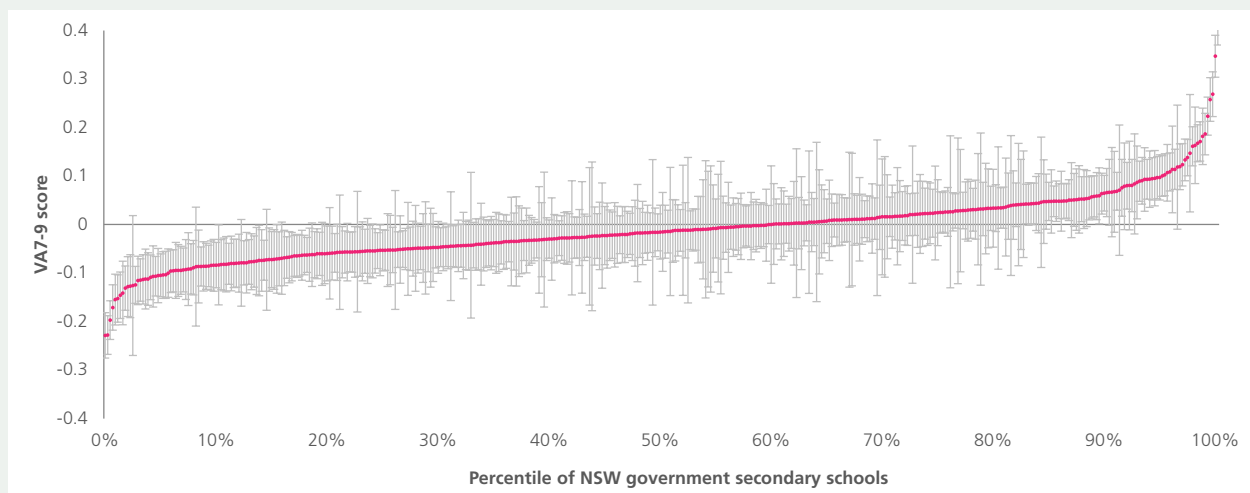
Reduced volatility

Our analysis of NSW government school data shows that gain score measures can be fairly unreliable and volatile for small schools, because they have fewer students. This makes it difficult to separate out the effect the school is having from the types of students that happened to be sitting the NAPLAN test at that school in a particular year. Without adjusting for this additional uncertainty, small schools are over-represented at the extreme ends of the gain distribution. In fact, one in five small schools switch from having significantly above average gain in one year to significantly below average gain in the next year, or vice versa. To reduce the volatility in VA estimates, particularly for small schools, the estimates are adjusted in proportion to their reliability. This means that estimates that we are less sure about are adjusted to be closer to the average school. This is important in NSW, which has a large number of small schools.

What do VA scores tell us about schools in NSW?

Figure 4 shows the distribution of VA7-9 scores with confidence intervals, across all NSW mainstream government schools with secondary students.

Figure 4 – Distribution of value-added (Years 7 to 9) scores across NSW government schools



The line at zero indicates the effect of the average school. The pattern shown in the figure is typical of what was found for all measures, and indicates that the majority of schools have overlapping confidence intervals, which straddle the average. This means that their effectiveness cannot be separated from the average school.

Around 35 per cent of schools have confidence intervals that are either entirely above or below the average line at zero. We can therefore be confident that their effectiveness is genuinely above or below that of the average school enrolling similar students.

The differences in contributions to learning between schools with high VA scores and schools with low VA scores can be substantial. One way to interpret these differences is in terms of the additional learning time achieved by a high VA school relative to a low VA school between two test times⁶. In terms of learning time, an average Year 7 student attending a 90th percentile school is around seven months ahead of a similar student attending a 10th percentile school by the time they reach Year 9 (Figure 5). For Years 3 to 5, the difference between a primary school at the 10th percentile and a primary school at the 90th percentile is about five months.

These results can inform how well school improvement strategies are working, and highlight possible areas for learning, future planning or further research. They can also provide low VA schools with learning opportunities by enabling investigation of the factors that contribute to higher VA scores in schools with similar student and school characteristics.

As for any single measure of school performance, VA scores are not perfect or definitive. The proposed VA measures will be most useful when they are used in conjunction with other measures to provide a profile of school effectiveness. VA scores also do not tell us *why* some schools add more value than others. However, they can be used as a starting point for in-depth analysis of schools that perform significantly better than others, to determine 'what works' to improve educational outcomes for students.

What does VA analysis tell us about which factors influence student outcomes?

Several factors are particularly important when predicting NAPLAN and HSC performance.

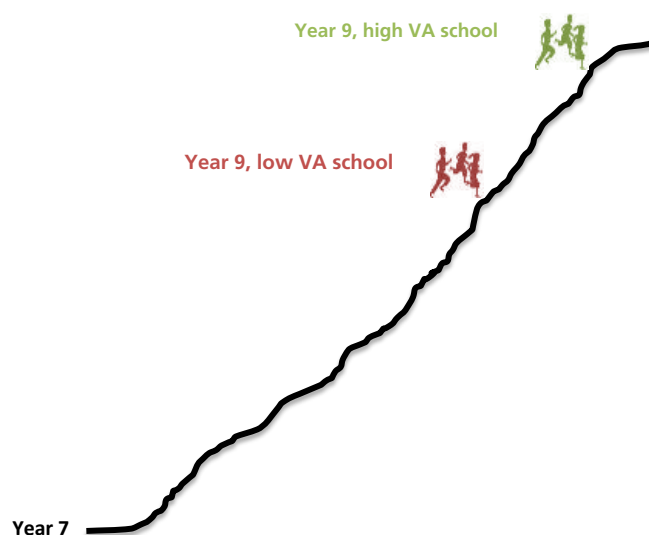
Prior ability

Of all of the factors, NAPLAN scores from prior years have the greatest influence on students' later scores. This is because a great deal of other information about the student, including their background, their cognitive ability, their attitudes, and their previous experiences at school, is embedded in prior achievement.

Student background factors

Prior achievement captures much of the influence of student level contextual factors such as whether they have an Aboriginal or Torres Strait Islander background and parental socio-economic status. The *additional* influence of the individual contextual factors on students' learning between two points in time is relatively small. However, these factors can accumulate

Figure 5 – The difference between a low VA school and a high VA school translates to up to seven months of learning time



Note: In this figure, low VA and high VA schools are defined as schools with VA scores in the 10th and 90th percentiles, respectively.

and become a significant barrier to achievement when a student experiences multiple aspects of disadvantage.

Clustering of high-performing students

Opportunity Classes and fully selective schools are two characteristics of the NSW education system that lead to many high-performing students all attending the same schools. The effect of selecting high-achieving students in a class or in a school is a persistent factor positively influencing student outcomes, even after variation in student prior achievement scores and other background factors have been taken into account⁷.

Gender effect

Our work also examined the effect of gender on student outcomes in different learning stages. We find that, having adjusted for prior scores and important contextual factors, the gender effect is negligible in all cases except in for the VA 9-12 (TES) measure. In senior secondary school, female students tend to experience more learning progress, mainly due to better results in English.



Single-sex schooling

The merit of educating students in single-sex schools as opposed to co-educational schools is a topic of some interest to the public and in academic research⁸. Overall, the evidence in this debate appears mixed⁹. It is not within the scope of this paper to comprehensively discuss all advantages and disadvantages of single-sex schooling for each gender. However, our VA analysis finds positive effects of single-sex schooling for both girls and boys in the NSW government system, after other student and school characteristics have been taken into account. This finding suggests further investigation is warranted as to reasons why students appear to have greater learning gains in these schools than in co-educational schools. In the interim, to ensure the fairness of the VA measures, we include a single-sex schooling factor in secondary VA models as the establishment of such schools is a system decision, and is out of the control of schools or principals.

Notes

1. For more information about the development of the VA measures, see the technical report prepared by the Centre for Education Statistics and Evaluation: CESE 2014, *Value added models for NSW schools*, prepared by L Lu and K Rickard, <http://www.cese.nsw.gov.au/component/k2/item/51-value-added-models-for-nsw-government-schools>
2. VA measures have been adopted by schooling systems in the United Kingdom, Hong Kong, more than 30 states or districts in the United States, and within Australia by the Victorian Department of Education and Early Childhood Development and the NSW Catholic education system. For more information on the research that underpins the VA methodology, see the technical report: CESE 2014, pp.6-21.
3. The Tertiary Entrance Score is a weighted score across the best ten HSC units a student attempted. This score is the basis for the calculation of the Australian Tertiary Admission Rank (ATAR).
4. VA measures are not suitable for schools for specific purposes, which enrol students with disabilities and special needs who may show different learning progress.
5. For more information about FOEI and how it was developed, see a summary at CESE 2013, *Getting the funding right*, Learning Curve 5, <http://www.cese.nsw.gov.au/component/k2/item/38-learning-curve-5>
6. Learning time is calculated with reference to the growth of the average student across two scholastic years. A student at a low VA school has growth from Year 7 to Year 9 that is about one-fifth lower than the growth of the average student. If the average student had 24 months of learning over this period, then a student at a low VA school has one-fifth less – about 19 months of equivalent learning.
7. It is worthwhile noting that the effect discussed here is the effect on individual students placed in an Opportunity Class or fully selective school. This is different from the impact that the streaming practice has on the system as a whole. The system-wide effect would also need to include the effect of such a practice on the other students who are not selected into selective classes or schools.
8. A Dabrowski 2014, *Single-sex schooling relies on myths of higher achievement*, The Conversation, 25 March 2014, <http://theconversation.com/single-sex-schooling-relies-on-myths-of-higher-achievement-23975>
9. U.S. Department of Education, Office of Planning, Evaluation and Policy Development 2005, *Single-sex versus coeducational schooling*, prepared by F Mael et al., pp.77-78.

Centre for Education Statistics and Evaluation
Office of Education
GPO Box 33
Sydney NSW 2001
Australia
T 9561 1211
E cese@det.nsw.edu.au
www.cese.nsw.gov.au

NSW Department of Education and Communities