

# Technical Notes for Tasmania's Education Performance Reports

## ***Part 1. The scope of the measures***

### **Why these measures?**

These measures include all the educationally significant data items for which the department collects valid and reliable data. They represent the broad range of measures of educational performance that jurisdictions seek to improve. Over time, these have proved to be reliable measures that are measured in ways consistent with other states and territories and reference national and international research. Because they have been collected for a number of years there is evidence that the measures are responsive to change, enabling the department to identify areas of concern or progress.

### **Why aren't other measures included, such as suspensions?**

Other measures are not included because they are either unavailable statewide on a recurrent basis, not as valid, or not as reliable. Student suspension measures, for example, are affected not only by student behaviour, but by the enforcement of disciplinary sanctions by principals: hence this measure is monitored, but not directly linked to Tasmania's educational performance. Over time, suspension measures have not been found to correlate as strongly and consistently with outcome measures as other data.

Reading has been chosen as the dominant measure of literacy, rather than writing, because it is a more reliable measure and is a stronger predictor of overall success in literacy. This is supported by international studies.

### **Which measures are more straightforward?**

Measures are reported as transparently as possible, however in some cases there are important technical considerations included in the measure. Here are some examples.

- Attendance rates are readily understandable as the proportion of students or staff attending daily through the year; technicalities only arise when considering rarer issues, such as students who leave the system without notification, determining when their enrolment ceases.
- Percentages of students achieving expected outcomes are readily understandable. There are, however, technical issues in determining these percentages, such as how students are being assessed in relation to the expected outcomes, and how the concept of expected outcome is defined, which itself is not trivial to determine for broad outcomes such as literacy and numeracy.
- Indexes of satisfaction or equity are not straightforward measures. There are some considerations.
- Satisfaction indexes treat survey responses across a range of questions, and use responses where the scale of agreement has involved a matter of degree, such as strongly disagree through to strongly agree.
- Similarly complex is the index of equity which involves measuring both student achievement and socioeconomic status based on separate data that must be matched to the student, and then computing indexes via a statistical method of regression analysis.

**Will the measures be reviewed and if so when?**

Yes, annually. The measures reported for 2008 have supplemented by two additional measures based on data from the National Assessment Program. Over time, it may be necessary to add new agreed measures to the priority areas which may supersede existing measures. If new measures become available that are reliable and valid, these will be additional to the current report. Retention is an area that needs attention because current measures are limited.

**Why are values considered Acceptable when the achievement is Low?**

The reporting of both achievement and improvement categories supports accountability while also recognising success, by acknowledging both the achievement in relation to absolute targets to reach, as well as monitoring evidence of improvement and sustaining achievement. Therefore, if the achievement is Low but there is evidence of improvement because achievement is trending up, the overall rating is Acceptable.

**Why are values different across measures that appear to be similar?**

There are different ways of measuring the same attribute. Measurement of heights, for example, may be measured in different units (centimetres, metres, feet and inches, etc.) depending on the ruler used. In the same way, measures of achievement in literacy and numeracy, or of satisfaction, depend on the measurement instrument being used. The figures reported are based on the best available measurement instruments and methods.

Satisfaction measures for staff, parents, and students are based on different sets of issues appropriate to their stakes in education. One cannot assume similar response patterns from these different groups, which is why each is surveyed, and is why reference to broader national data for each is required to assist interpretation of each value as being High, Low or Intermediate.

***Part 2. Measurement issues in evaluation*****How are the achievement ranges defined to determine achievement?**

The intermediate ranges have been determined by referencing against national data where they are available. In some cases where complete national data are unavailable, related data have been used to estimate a national value.

The intermediate ranges have been defined by considering the spread or deviation of the data available, including Tasmanian data over time. In addition, where the available data have been estimated or have a less established historical base, slightly wider boundaries have been provided due to the reduced precision.

**How are the improvement categories determined?**

Improvements are determined based on the set of values available for the current year and the previous three years. In determining improvement categories, improvement estimates (or trends) are calculated by balancing both short-term change, such as 2006 to 2007, with longer-term change, such as 2005 to 2008. In some cases, an exceptional high or low value in 2007 would yield quite different improvement estimates according to whether one took a short-term or long-term view of change. In other instances, older data, such as 2005, do not reflect the real changes that might currently be happening. In determining the right balance, slope estimates from regression line are determined for 2005–2008, 2006–2008, and 2007–2008 and then averaged: this effectively gives slightly greater balance to short-term change, moderated against longer-term change.

Once the improvement estimates have been determined, a judgement is made regarding how slight changes, such as 0.5% per year, may be considered as “Stable”, “Trend Up”, or “Trend Down”. This decision differs for each measure, depending upon a review of the stability of the measure over time, a review of the stability of the trend estimates, and degree of precision in the available data.

**Will the achievement ranges and improvement estimates be reviewed and if so when?**

Only if national data not currently available comes to hand or, if in later years, new improvement targets need to be set.

**Part 3. Referencing External Data for Determining Boundaries**

**Readiness for school: Percentage of Kindergarten students achieving expected outcomes**

These data are referenced to Tasmanian historical data, as this assessment instrument was developed in Tasmania.

Further information, including the assessment instrument and support materials for teaching to address various criteria, is available from

<http://www.education.tas.gov.au/school/educators/curriculum/kindergartencheck>.

**Early literacy and numeracy: Percentage of Prep students achieving expected outcome**

The Performance Indicators in Primary Schools (PIPS) test was developed at Durham University in the UK, where it is widely used. It is used in several Australian states and territories, with administration in Australia licensed via the University of Western Australia. <http://www.education.uwa.edu.au/pips>.

Students’ results are scored for reading from 0–193, and for mathematics from 0–69. Tasmania has historically reported to schools each year the percentages of students one standard deviation below the state mean score, that is generally the lowest 17–18% of student results statewide. For a system-report, however, this would not monitor change over time—the achievement value each year would be constant around 82%. A review of the data led to determining this approximate level of the expected outcomes to be a set score of 66 for reading, and a score of 41 for mathematics, set irrespective of calendar year. The intermediate range of 81–83 was informed by Tasmanian history 2004–2006 with 82% achieving outcomes at or above these set scores.

There are no known published data concerning PIPS performance across Australia in a comparative form using these set scores. Yet to be published research reports indicate, from sample data reported along a logit scale (a technical measure of average students performance), that Tasmanian performance is at or slightly above the average score of other states using the PIPS assessment. These preliminary data could have been used to propose a lower national reference estimate, such as 81 with intermediate range 80–82, however the range of 81–83 is used to set a high yet realistic expectation.

In the national context, a broader assessment of social indicators is to be implemented in 2009, the Australian Early Development Index (AEDI)

[http://www.rch.org.au/australianedi/index.cfm?doc\\_id=6210](http://www.rch.org.au/australianedi/index.cfm?doc_id=6210). Depending on its validity and reliability as well as its application to measuring educational performance, the AEDI may be added to the performance report in later years.

### **Literacy and numeracy testing: Percentage of Years 3, 5, 7 and 9 students achieving expected outcomes**

These percentages are determined in relation to the percentages of students achieving at or above the national minimum standard in reading and numeracy from the NAPLAN assessments. Further information, including official results for each state across all domains, and for various year levels and student groups, are available from [www.naplan.edu.au](http://www.naplan.edu.au).

The figures shown are based on the official Tasmanian figures of students achieving the national minimum standard, averaged across years 3, 5, 7 and 9. These figures are the best measure available, as they address technical statistical considerations related to test construction and absent students.

The intermediate range was developed from the official Australian figures of students achieving the national minimum standard, averaged across years 3, 5, 7 and 9.

### **Literacy and numeracy testing: Index of gain for Years 3–5, 5–7, 7–9**

These gains are improvement in test scores over two-year intervals. The measure averages gains from Years 3–5, 5–7 and 7–9. In general, gains on this measurement scale appear greater from Years 3–5 than from Years 5–7 or from Years 7–9. The estimates may also be affected by selective absenteeism of students in one test: these students' results are excluded from analysis as they have no measure of gain available.

While test scores for 2008 are nationally comparable on the NAPLAN scale, gains in test scores rely also on the test scores from 2006, which were based on various historical scales within each state. A process of equating has been used to place student NAPLAN scores from 2008 on to the historical Tasmania scale, and hence determine the gain in each student's test score.

From 2010 it will be possible to have nationally comparable data on student gains using NAPLAN scores. In the absence of a specific national value, the range of 20–25 has been based on national evidence from NAPLAN (2008) and benchmark results (2007 and previous). This range sets a high yet realistic expectation without overstating Tasmania's outcomes in a national context.

### **Student attendance: Rate of student attendance**

Complete national data are yet to be published, however some states have published performance data related to student absence rates.

These attendance rates have informed the intermediate range of 91–93. In general, Tasmanian data have undergone processes not employed by some states, such as adjustments for school closures on selected days. This adjustment has the effect of decreasing the Tasmanian attendance rates.

### **Student retention: Rate of students retained Year 10 to Year 12 (apparent/direct)**

National data are published each year for apparent retention rates on Australian Bureau of Statistics website, Schools Australia.

In general, Tasmanian apparent retention rates have been about 5% below the national reference for government schools (~70), and 10% below the national reference for all schools (~75). The intermediate range (70–80) has used the higher value as the target for Tasmania.

The methodology used for apparent retention nationally cannot be used at the individual school level in Tasmania because of Tasmania's college system, where students do not stay at the same high school until the end of year 12. Direct retention rates are therefore used for the regional summary and these are not published nationally. Tasmanian historical data are 45–49, however this report has used the known 10% difference to the national apparent retention rate to set the intermediate range (55–59) at a higher value as the target for Tasmania.

### **School effectiveness: Index of school effectiveness and improvement**

These data are referenced to Tasmanian data in 2006, as this measure was developed in Tasmania. The index is an aggregation of a range of available measures used to indicate aspects of school effectiveness and improvement. It is a complex measure set on a scale of 0–100. A score of 100 would indicate that all schools had generally improved and performed at expected levels. The intermediate range has been set higher than the 2006 figure, however, to set a challenging yet realistic target range for future years.

### **Staff satisfaction: Index of staff general satisfaction**

All staff are invited to respond to an online survey of Organisational Health, developed by an independent contractor, InsightSRC, which is used in education workplaces as well as broader government and private sectors around Australia. The survey includes various modules measuring aspects such as motivation, empathy, clarity, engagement, appraisal and recognition, job satisfaction, curriculum coordination, student orientation, and excessive work demands.

The measure of staff satisfaction is an index across the set of questions asked, known as the staff climate index. The scale reported is from 0–10, the original scale was 0–100, however because the measure is NOT a percentage of staff satisfied, but an index of the average degree to which they are satisfied, the reported value is rescaled from 0–10 to avoid confusion. National benchmark data for Australian education P–12 are prepared by InsightSRC.

### **Parent satisfaction: Percentage of parents generally satisfied**

The parent satisfaction survey is used with the permission of the Victorian Government.

National benchmark data for Australian education P–12 are prepared by InsightSRC. The values indicate the average degree to which parents are satisfied.

The percentage reported in *Tasmania's Education Performance Report 2008* does not use the scale of average degree to which they are satisfied, but is based upon the percentages of parents who reported general satisfaction for a single summary question of overall satisfaction. This has been used to ensure continuity with department annual reports in recent years 2003–2005 and ensures continuity into the future, based on the question that is the same each year, despite changes to the set of questions on the parent survey over years.

### **Parent satisfaction: Index of parent satisfaction with reporting**

The parent satisfaction survey is used with the permission of the Victorian Government. National benchmark data for Australian education P–12 are prepared by InsightSRC. The values indicate the average degree to which parents are satisfied.

The index reported in Tasmania's Education Performance Reports does not use the scale of average degree to which they are satisfied, but is based upon the percentages of parents who reported general satisfaction with reporting across three questions.

### **Student satisfaction: Index of student general satisfaction**

The student satisfaction survey is used with the permission of the Victorian Government. Students Year 5 and above, randomly selected from each school and college, respond to an online survey, developed for the Victorian Government, which is used in other schools around Australia. The survey includes various modules measuring aspects such as morale, aspects of teaching and learning, and student relationships.

National benchmark data for Australian education P–12, on a scale of 0–100, are prepared by InsightSRC. The measure of student satisfaction is an index across the set of questions related to teaching and learning, known as the student climate index. The scale reported is from 0–10, the original scale was 0–100, however, because the measure is NOT a percentage of students satisfied, but an index of the average degree to which they are satisfied; the reported value is rescaled as 0–10 to avoid confusion.

### **Staff attendance: Rate of staff attendance**

A report on the Australian public service served as the basis for national referencing. <http://www.apsc.gov.au/stateoftheservice/0607/downloads.htm>.

This report found an average value of 9.4 days absent per annum. If one estimates approximately 241 working days as expected per annum, accounting for weekends, public holidays and annual recreation leave, the absence estimate is  $9.4/241 = 3.9\%$ , hence a national attendance estimate of 96.1%.

### **Indigenous equity: Percentage gap in students achieving expected outcomes**

The percentages of students achieving at or above the national minimum standards for reading and numeracy in Years 3, 5, 7 and 9 are reported nationally for all students and for Indigenous students. To establish the intermediate range, gaps between all students and Indigenous students were averaged for the national data.

### **Socioeconomic equity: Index of equity of achievement by socioeconomic status**

The Program of International Student Achievement (PISA), is an international education project, with significant involvement of the Australian Council for Educational Research (ACER). Further information about PISA, including Australian major findings, is available from [www.acer.edu.au](http://www.acer.edu.au).

One published outcome for PISA 2006 is that socioeconomic inequity, as measured by the slope of the regression line (the effect of socioeconomic status on student performance in scientific literacy), is greater in Tasmania than for the nation.

The Tasmanian data measuring slope is based on approximately 18,000 students across Years 3, 5, 7 and 9 for both reading and numeracy performance. Socioeconomic status for each student has been based on the highest educational level reported by the parents of the student. When these national results are aligned to the Tasmanian scale scores, they support an intermediate range of 8–10.