MEASURING LEARNING GAIN IN NAPLAN

QUALIFICATIONS CONCERNING NAPLAN

The National Assessment Program – Literacy and Numeracy (NAPLAN) is not an ideal instrument for investigating the learning culture of a school for many reasons.

The original intention of NAPLAN was diagnostic. As the test has become increasingly high-stakes (for example through publication of data on MySchool) some schools have instituted practices to boost scores while not addressing the underlying constructs of numeracy and literacy that should be the diagnostic focus of any school.

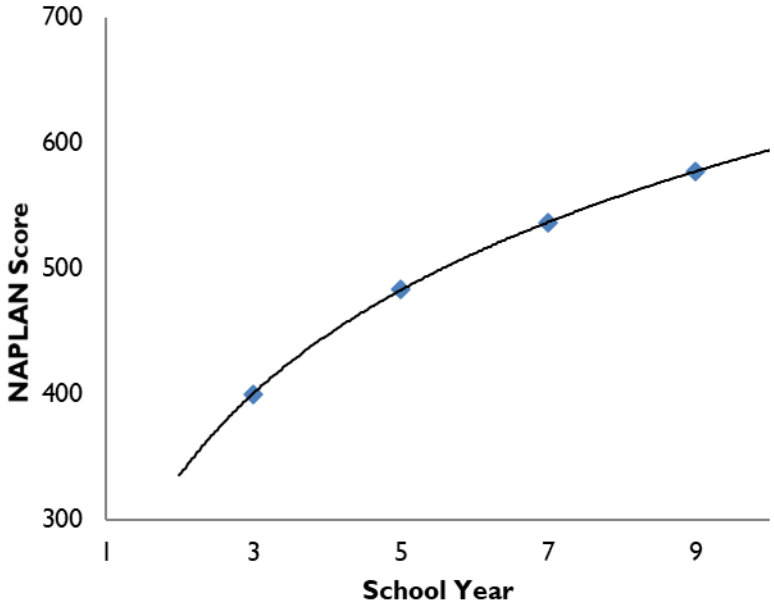
1. The proper aims of a learning culture are far broader than simply literacy and numeracy. Student achievement is based on engagement, wellbeing, future-focus, resilience and self-concept. While these do impact NAPLAN results, some direct measurement of such underpinnings is needed for a fuller picture of a school's learning culture.
2. NAPLAN does not distinguish well between high-achieving students. For high-achieving students other data would be needed to assess the impact of the learning culture. As a "basic skills" test, there is a ceiling in tests like NAPLAN approaching which able students cannot adequately demonstrate their capabilities.

THE ISSUE WITH MEASURING LEARNING GAIN

Not all learning gain scores in NAPLAN are equal. The size of the learning gain depends on the starting point on the scale. For example, the average Year 5 to Year 7 learning gain in Reading in NSW is always larger than that for Year 7 to Year 9. Comparisons of learning gain figures therefore are fraught. If one Year 7 class begins from a mean score of 540 but another begins from a mean of 620, the first class typically will have a much larger learning gain than the second, even if both have made average two-year progress from their starting point. Therefore, comparison of each to the mean learning gain for all of Year 7 gives an invalid picture.

The problem arises from the nature of the NAPLAN scale, as shown in Fig. 1. If the growth curve were a straight line, there would be no problem. Students would be expected to make equal increases on the scale in equal periods of time, no matter what point they started at. The problem arises from the curvilinear nature of the growth path.

Fig. 1: NAPLAN Growth Curve



NAPLAN scores are constructed using a Rasch model. The scale is common both across all academic years and also across successive calendar years (an effect achieved by item equalization). Because growth in literacy (as measured on such a scale) is constrained, the growth curve is logarithmic as shown.

What is needed therefore is a means of placing measures of learning gain on a common scale, and then determining the extent to which a particular class or individual has made 'typical' progress, independently of the class's or student's starting point.

The NAPLAN growth curve was defined by the parameters assigned in 2008, the first year of the tests. Subsequent years have been longitudinally aligned to the original dataset.

Given the following terms

y	Academic year of the students (3, 5, 7, 9)
y_e	Decimal-equivalent year-value of a particular score on the NAPLAN scale
α_y	Score on a NAPLAN scale in academic year *y*
α_{y+2}	Score on a NAPLAN scale two years later, academic year *y+2*
$\hat{\alpha}$	Typical score on a NAPLAN scale in year y+2
β	Coefficient of the natural logarithm of *y*
χ	Gain Constant
$\hat{\Delta}$	Typical difference (gain score) between years *y* and *y+2*

the growth curve in Fig. 1 has the general form

$$\alpha_y = \beta \ln y + \chi \tag{I}$$

Mapping to the original defining NAPLAN data gives

$$\beta = 161.19$$

$$\chi = 223.79$$

with

$$R^2 = 0.9999$$

It follows from equation (1) that

$$y_e = e^{\left(\frac{\alpha-\chi}{\beta}\right)} \tag{2}$$

Combining equations (1) and (2) then enables the calculation of the Typical Gain over two years from any score α_y as

$$\hat{\Delta} = \beta \ln\left(e^{\left(\frac{\alpha_y-\chi}{\beta}\right)} + 2 \right) + \chi - \alpha_y \tag{3}$$

It is of particular importance that equation (3) has no dependence on the academic year; it is simply a model based on the starting score α_y.

Equation (3) therefore enables for any class in any academic year in any test domain the calculation of a Typical Mean Gain, and comparison of the Raw Mean Gain with this. If the Raw Mean Gain equals $\hat{\Delta}$ then the cohort has made two years of learning gain over the passage of the two years between the tests. If it is greater, then the learning gain is proportionally more. Likewise, if it is less, then the learning gain is proportionally less.