

## 9. NCEA results and our competency measures

The NCEA is a new way of measuring student achievement, and has proved to be somewhat controversial, with many claims and counterclaims about how understandable the results are, how well they measure student ability, how robust the system is, and how well it serves to predict success at university.

The Competent Children, Competent Learners study provided us with an opportunity to measure the students' NCEA results against our cognitive and attitudinal competency measures. We have separate literacy, numeracy, and logical problem-solving measures, but for the comparison it is probably most helpful to use the age-16 composite competency measure, as this measure covers all three aspects together, and to look at the overall pattern of achievement in the NCEA (rather than looking at literacy and/or numeracy standards separately). This allows us to compare our measures and those from NCEA overall as measures of cognitive achievement.

What comparisons are meaningful? There are two fundamental questions we can ask of the data. Firstly, how well do our measures of achievement and engagement predict success in the NCEA? And secondly, the reverse question: How well do NCEA results (overall) predict literacy and numeracy as measured by our composite cognitive competency?

We look first at the possible measures of achievement derived from NCEA results, then at the inter-correlations between the possible measures, and lastly use linear models to investigate the two comparability questions.

### NCEA measures

The measures we have used of achievement in the NCEA are:

- the percentage of achievement standards attempted that were achieved (the number of As divided by the total number of A, M, E, and N results for achievement standards<sup>1</sup>)
- the percentage of achievement standards attempted that were merits (Ms)
- the percentage of achievement standards that were excellences (Es)
- the total number of credits (Unit Standards or Achievement Standards) gained
- the total number of Level 1 credits achieved.

These measures are all simple to calculate from NCEA results, and do not attempt to discriminate in any way between the actual standards attempted. There is no attempt to distinguish qualitatively between possible standards, nor is there any attempt to distinguish between courses that "measure" basic literacy or numeracy and those that "measure" other skills.

### Correlations

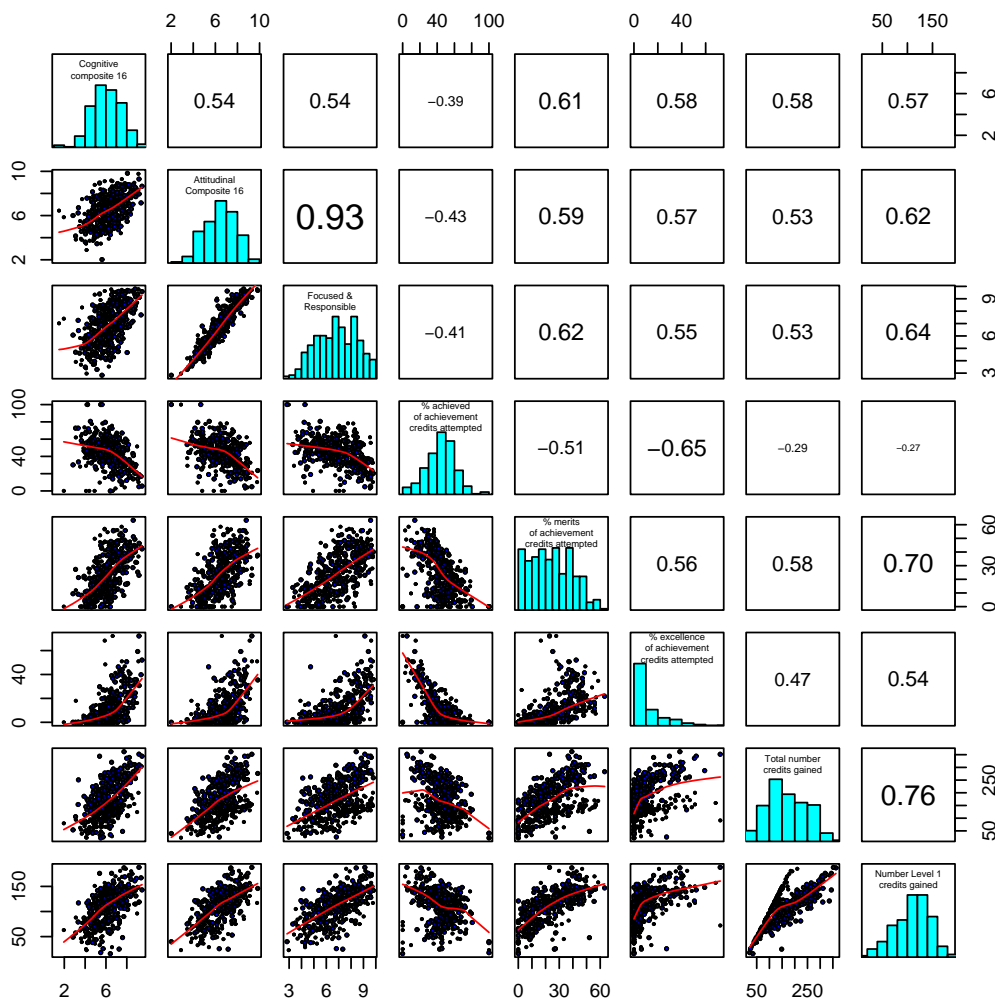
Preliminary investigations showed that the NCEA measures were more strongly associated with *focused and responsible* than with *attitudinal composite* or any of the other attitudinal competencies, in particular the two social skills measures. As far as the cognitive competencies are concerned, the *cognitive composite* was more strongly associated with the NCEA measures than were any of the three separate competencies. Figure 2 shows

---

<sup>1</sup> A=Achieved, M=Merit, E=Excellence, and N=Not achieved.

the distribution of the achievement measures down the diagonal (the little histograms); the scatter plots showing the relationship between the measures, and the corresponding correlations.

Figure 2: Correlations between cognitive and attitudinal competencies and NCEA measures



The first three rows and columns show the Competent Learners competencies, and the bottom five rows and right-hand five columns show the data for the NCEA results. The strongest correlations are between the three Competent Learners competencies (particularly between *attitudinal composite* and *focused and responsible*, one of the competencies used to construct the composite), and between the five NCEA measures, particularly between the number of Level 1 credits and the total number of credits gained, and the percentage of merits gained out of the number of achievement standards attempted. The percentage achievement standards getting an A (Achieved) is negatively correlated with the other variables, as getting a high percentage of just achieved is typically not an indication of academic excellence.

The lines through the scatter plots give an indication of the extent to which the typical assumption that the relationships measured are linear (if the line is straight) is satisfied. The percentage of Es gives several indications of departure from linearity, mainly because most students received very few Es, but a few received a high percentage of Es.

The total number of credits gained is heavily dependent on the year (we had students in both Year 11 and Year 12). In a plot where the year levels are distinguished, the points corresponding to Year 11 students lie mainly under the slightly wiggly line, and those for Year 12 lie mainly above it in each of the plots involving the total number of credits gained. In the plots against *attitudinal composite*, *focused and responsible*, and the percentage of merits achieved it is actually possible to make out two separate clouds of points, running at slightly different angles: the Year 11s below the wiggly line (with a less steep gradient), and the Year 12s above (with a steeper gradient). This separation is clearest of all for the total number of credits plotted against the number of Level 1 credits, as Year 12 students have more credits in total, but for most Year 11 students the two measures are the same or very similar.

In education, correlations in the region of 0.5–0.6 represent relatively strong associations. For example, the correlations between our *literacy*, *numeracy*, and *logical problem-solving* scores are between 0.51 (between *literacy* and *logical problem-solving*) and 0.68 (between *numeracy* and *logical problem-solving*). The correlations between the NCEA measures and our competency measures are all between 0.43 and 0.60 (ignoring sign); in other words, they are all relatively strong. As the correlations between *focused and responsible* and the NCEA variables are stronger than the corresponding measures using *attitudinal composite*, we use *focused and responsible* in the models fitted.

If associations exist, how well can we predict the one set of outcomes from the other?

## Models

Each of the models that follow is based on different numbers of observations. We had a total of 447 students, 27 of whom had left school (so do not have age-16 attitudinal competency measures), five are not in mainstream school (again, no attitudinal competency measures), two or three could or would not complete the cognitive competency tasks, up to another 10 have one or more of the attitudinal competency measures missing as one or more of their teachers gave no response to too many of the items used to form the attitudinal competencies, up to seven were not assessed using NCEA, and about four students refused us permission to access their NCEA results.

The numbers of students excluded from each of the models is given in Table 85.

**Table 85: Number of students with missing data**

Model	Year 11	Year 12	Total
Total number of credits	40	47	80
Percent of A results			52
Percent of M results			64
Percent of E results			64
Cognitive competency			46

Imputing missing values is increasingly popular and commonplace. However, most of the missing values in these data are missing because of a difference between the individuals with measures and those without them. Imputation would make these students “the same” as the rest, as the imputations could not be made taking the differences into account, as it is these differences that were the initial cause for the missingness in the first place.

A consequence of excluding these students from the modelling process is that the models can only be regarded as applying to students in mainstream schools, doing the kinds of subjects and qualifications that do not give rise to missing data in any of the ways described above.

### Predicting NCEA results from Competent Learners measures

We look in turn at predicting the *number of credits gained*, the *number of Level 1 credits gained*, the *percentage of just achieved*, the *percentage of merits achieved*, and the *percentage of excellences achieved*. The original model fitted included our *cognitive composite*, *focused and responsible*, year level (and we tested for interactions between year level and the other variables, in case the effect of some was different as the students neared the end of their secondary education), the English teachers' judgement of the students' attitudes to NCEA and working for it (*NCEA assessment*), and our three school-wide engagement variables: *engaged at school*, *confident at school*, and *satisfaction with subject mix*.

#### Total number of credits gained

It was necessary to model the results for Year 11 ( $n = 155$ ) and Year 12 ( $n = 261$ ) students separately.<sup>2</sup> This allows different gradients (different parameter estimates) for Year 11 and Year 12 students, as suggested by the scatter plots in the bottom row in Figure 2. This difference in gradient is possibly because Year 12 students, particularly the more able ones, do more high-credit standards than Year 11 students. The total credit score of Year 12 students increases more rapidly with increasing ability than the score for Year 11 students does.

We could account for 68 percent of the variability in the total number of credits gained by Year 11 students with a model that included *focused and responsible*, *NCEA assessment* (English teachers' views of the students in relation to NCEA), and *cognitive composite*. The three school-wide engagement measures did not add significantly to the models (and in fact were too strongly correlated with the variables already in the model to be included, to avoid multicollinearity).

**Table 86: Model to predict the total number of credits gained for Year 11 students**

Parameter	Estimate	Standard error	p-value	LMG (%) (confidence interval)	Partial correlation
Intercept	-50.49	10.01	< 0.0001		
Cognitive composite	14.28	1.88	< 0.0001	28.6 (22.9, 34.9)	0.56
Focused and responsible	6.89	1.95	0.0006	23.0 (18.7, 27.8)	0.29
NCEA assessment	9.80	3.50	0.006	15.9 (11.5, 20.7)	0.23

The picture for Year 12 students is very similar; the biggest difference is in the rate of increase in the number of credits achieved for each unit increase in each of the explanatory variables. This model accounted for 60 percent of the variability in the total number of credits gained.

<sup>2</sup> A single model, with year level included in an interaction term (required for the differing gradients, particularly for *focused and responsible*, at the two year levels) provided parameter estimates that suggested the data were collinear (in particular, negative parameter estimates for some of the terms that clearly have a positive relationship, and very large standard errors for those estimates). The estimates for models for the years fitted separately, and the variance inflation factors for the models, suggest that the separate models give an adequate and consistent description of the data.

**Table 87: Model to predict the total number of credits gained for Year 12 students**

Parameter	Estimate	Standard error	p-value	LMG (%) (confidence interval)	Partial correlation
Intercept	-74.26	15.63	< 0.0001		
Cognitive composite	17.80	2.46	< 0.0001	20.9 (16.1, 26.6)	0.42
Focused and responsible	16.50	2.79	< 0.0001	24.4 (20.1, 29.3)	0.36
NCEA assessment	15.87	5.28	0.003	14.7 (10.2, 20.0)	0.19

What we can read from the models (in a sense, working backwards) is that the total number of credits gained is likely to be a useful measure at either year level of someone's overall cognitive ability. Higher numbers of credits gained are also associated with positive attitudes to work, in particular work for the NCEA, and the ability to focus on the task in hand and take responsibility.

For illustrative purposes, we can divide the explanatory variables into quartile groups, and then calculate the average number of credits gained by students in each of these groups (Table 88). Notice that the number of credits achieved by the Year 12 students in our study is markedly more variable than the number achieved by Year 11 students (the standard deviation is in the order of twice as big).

**Table 88: Mean (and standard deviation) of total number of credits by students of different competency levels**

Quartile group	Cognitive composite		Focused and responsible		NCEA assessment	
	Year 11	Year 12	Year 11	Year 12	Year 11	Year 12
Lowest	78.7 (29.6)	146.2 (51.0)	79.2 (28.9)	142.8 (53.5)	88.7 (35.1)	149.9 (48.4)
Second lowest	111.4 (31.6)	179.9 (63.1)	99.6 (35.8)	186.0 (52.8)	107.4 (28.7)	185.8 (63.3)
Second highest	121.3 (31.4)	210.3 (60.0)	127.4 (29.5)	221.8 (51.9)	121.9 (35.9)	223.8 (51.5)
Highest	154.9 (26.5)	249.5 (60.0)	145.9 (24.7)	270.0 (43.3)	147.5 (27.0)	253.3 (60.3)

### *Number of Level 1 credits gained*

The situation is much the same as for the total number of credits gained. The model for Year 11 students accounted for 54 percent of the variability in number of Level 1 credits gained.

**Table 89: Model to predict the total number of Level 1 credits gained for Year 11 students**

Parameter	Estimate	Standard error	p-value	LMG (%) (confidence interval)	Partial correlation
Intercept	-27.9	10.6	0.010		
Cognitive composite	9.6	1.9	< 0.0001	22 (16, 30)	0.38
Focused and responsible	11.1	1.5	< 0.0001	32 (25, 40)	0.53

The picture for Year 12 students is very similar; the biggest difference is in the lower rate of increase in the number of credits achieved for each unit increase in each of the explanatory variables. This model accounted for 45 percent of the variability in the number of Level 1 credits gained.

Table 90: Model to predict the total number of Level 1 credits gained for Year 12 students

Parameter	Estimate	Standard error	p-value	LMG (%) (confidence interval)	Partial correlation
Intercept	9.3	7.6	0.222		
Cognitive composite	7.1	1.3	< 0.0001	19 (14, 25)	0.34
Focused and responsible	8.8	1.1	< 0.0001	27 (20,33)	0.46

Because some of the Year 12 students increased their number of Level 1 credits while in Year 12, the association between *cognitive competency* and *number of Level 1 NCEA credits* is weaker for Year 12 students; the number of Level 1 credits achieved **by Year 11** is a better indicator of cognitive competency and attitude than a more general total where there is no distinction between credits achieved in or after Year 11.

*Percentage of achievement standards that were "Achieved"*

As this is a measure of the *percentage* of standards achieved, not the total number of credits, there were no marked differences between Year 11 and Year 12 students. A single model could be fitted and this model accounted for 20 percent of the variability in the percent of A results (out of the number of A, M, E, and N results). The only explanatory variables that were statistically significant in the model were the two composite competency measures (Table 91).

Table 91: Model for percentage of A results

Parameter	Estimate	Standard error	p-value	LMG (%) (confidence interval)	Partial correlation
Intercept	83.2	3.97	< 0.0001		
Cognitive composite	-3.09	0.65	< 0.0001	10.2 (5.9, 16.2)	-0.24
Attitudinal composite	-3.06	0.63	< 0.0001	10.3 (6.5, 14.8)	-0.24

The relationship between the competency measures and the percentage of As is *negative*, as the students with *lower* competency scores tended to have *higher* proportions of A results. For each unit increase in the explanatory variables (on the 1–10 competency scales), on average, the students had about 3 percent fewer A results (and so correspondingly more M or E results).

A summary of the differences across competency levels, and lack of difference across year levels is given in Table 92. The results for the two year levels are very similar, although the lower percentages for the higher quartile groups are more marked for Year 12 students, and the results for the Year 12 students are consistently slightly less variable than those for the Year 11 students.

**Table 92: Mean (and standard deviation) of percentage of A achievement standard results by students of different competency levels**

Quartile group	Cognitive composite		Attitudinal composite	
	Year 11	Year 12	Year 11	Year 12
Lowest	49.3 (22.1)	51.3 (14.4)	51.8 (17.0)	51.8 (16.1)
Second lowest	48.8 (15.8)	50.2 (13.2)	46.8 (18.8)	48.5 (11.0)
Second highest	43.1 (16.8)	45.5 (14.1)	46.2 (18.5)	43.1 (12.0)
Highest	33.0 (17.4)	33.9 (14.2)	35.0 (17.3)	33.7 (14.6)

### *Percentage of achievement standards that were "Merit"*

For this measure there was a difference between Year 11 and Year 12 students. A single model could be fitted (the gradients were the same for the Year 11 and Year 12 students) and this model accounted for 53 percent of the variability in the percent of M results (out of the number of A, M, E, and N results). The explanatory variables that were statistically significant in the model were the *cognitive composite*, *NCEA assessment*, *focused and responsible*, *affirmed at school*, and year level (Table 93).

**Table 93: Model for percentage of M results**

Parameter	Estimate	Standard error	p-value	LMG (%) (confidence interval)	Partial correlation
Intercept	-11.75	13.19	0.37		
Cognitive composite	4.71	0.49	< 0.0001	21.4 (17.5, 26.2)	0.44
Focused and responsible	2.66	0.56	< 0.0001	17.2 (14.3, 20.3)	0.24
NCEA assessment	2.31	1.03	0.025	9.6 (7.2, 12.5)	0.11
Affirmed at school	3.00	1.37	0.029	4.0 (2.4, 6.3)	0.11
Year level (if Year 12)	-2.64	1.15	0.022	3.6 (0.1, 1.4)	-0.11

For each unit increase in the explanatory variables, other than year level (that is, for those on the 1–10 competency scales), on average, the students had 2 to 5 percent more M results. Year 12 students had 2 to 3 percent fewer M results than Year 11 students.

A summary of the differences across competency levels for the three explanatory variables that accounted for the most variance in the percentage of M results, and lack of difference across year levels is given in Table 94. The results for the two year levels are very similar, although the lower percentages for the higher quartile groups are more marked for Year 12 students, and the results for the Year 12 students are consistently slightly less variable than those for the Year 11 students.

**Table 94: Mean (and standard deviation) of percentage of M achievement standard results by students of different competency levels**

Quartile group	Cognitive composite		Focused and responsible		NCEA assessment	
	Year 11	Year 12	Year 11	Year 12	Year 11	Year 12
Lowest	11.0 (12.3)	10.8 (8.5)	11.4 (11.3)	11.6 (9.4)	17.1 (15.2)	13.4 (11.0)
Second lowest	23.6 (15.3)	18.8 (10.4)	24.2 (13.5)	20.8 (10.6)	21.4 (13.5)	21.0 (13.2)
Second highest	29.4 (13.8)	25.7 (14.0)	28.2 (17.6)	28.8 (13.6)	29.0 (17.1)	27.7 (12.3)
Highest	38.8 (10.4)	35.6 (10.9)	34.0 (13.1)	36.9 (10.2)	33.9 (14.1)	33.9 (12.2)

There is a clear trend of increasing percentage of M results with increasing competency measure, and for Year 11 students to have achieved a slightly higher percentage of M results than Year 12 students. The less able students achieve Merit in just over 10 percent of their achievement standards, compared with the most able students who on average have Merit in over a third of their achievement standards.

*Percentage of achievement standards that were “Excellence”*

For this measure there was a difference between Year 11 and Year 12 students. A single model could be fitted (the gradients were the same for the Year 11 and Year 12 students) and this model accounted for 52 percent of the variability in the percent of E results (out of the number of A, M, E, and N results). The explanatory variables that were statistically significant in the model were the *cognitive composite*, *attitudinal composite*, *NCEA assessment*, and year level (Table 95).

**Table 95: Model for percentage of E results**

Parameter	Estimate	Standard error	p-value	LMG (%) (confidence interval)	Partial correlation
Intercept	-12.07	11.91	0.31		
Cognitive composite	4.32	0.44	< 0.0001	23 (17, 26)	0.44
Attitudinal composite	2.36	0.55	< 0.0001	15 (12, 19)	0.22
NCEA assessment	2.29	0.87	0.009	9 (6, 12)	0.13
Year level (if Year 12)	-2.32	1.02	0.024	0.4 (0.1, 1)	-0.11

For each unit increase in the explanatory variables, other than year level (those on the 1–10 competency scales), on average, the students had 2 to 4 percent more E results. Year 12 students had about 2 percent fewer E results than Year 11 students.

A summary of the differences across competency levels for the three explanatory variables that accounted for the most variance in the percentage of E results, and lack of difference across year levels is given in Table 96. The results for the two year levels are similar, although the lower percentages for the higher quartile groups are more marked for Year 12 students, and the results for the Year 12 students are consistently slightly less variable than those for the Year 11 students.

**Table 96: Mean (and standard deviation) of percentage of E achievement standard results by students of different competency levels**

Quartile group	Cognitive composite		Attitudinal composite		NCEA assessment	
	Year 11	Year 12	Year 11	Year 12	Year 11	Year 12
Lowest	2.2 (4.2)	2.0 (3.2)	2.4 (4.6)	2.8 (6.7)	5.7 (9.4)	3.0 (5.1)
Second lowest	9.4 (13.2)	5.3 (6.7)	8.4 (9.7)	6.5 (7.5)	9.9 (14.0)	6.7 (9.1)
Second highest	13.3 (15.1)	8.8 (10.0)	11.0 (14.0)	10.7 (11.0)	12.4 (13.1)	10.3 (10.7)
Highest	24.9 (14.7)	21.3 (15.8)	24.1 (16.7)	22.2 (16.2)	19.9 (17.9)	21.7 (16.8)

Students in the lowest quartile groups tended to achieve Excellence for only about 2 percent of their achievement standards, while those in the highest quartile groups tended to achieve Excellence for a fifth to a quarter of their achievement standards. The variability in the percentage of E results increased with increasing ability level.<sup>3</sup>

### Predicting Competent Learners measures from NCEA results

If knowing something about attitudinal and cognitive competencies allows relatively good predictions (for education) of NCEA results, does the process work in reverse? If we know NCEA results, can we predict cognitive competency?

A model including the total number of credits gained, the percentage of the achievement standards (AS) that were Merit, and the percentage that were Excellence accounted for 51 percent of the variability in *cognitive competency*.

**Table 97: Model for cognitive composite predicted by NCEA results**

Parameter	Estimate	Standard error	p-value	LMG (%) (confidence interval)	Partial correlation
Intercept	4.488	0.120	< 0.0001		
Percent AS that were Excellence	0.032	0.0046	< 0.0001	18 (14, 22)	0.33
Percent AS that were Merit	0.024	0.0041	< 0.0001	17 (14, 21)	0.27
Total number credits gained	0.0047	0.0008	< 0.0001	16 (12, 20)	0.29

## Conclusion

Our competency measures are measures of attitudinal and social competency, and of literacy, numeracy, and problem solving. NCEA results are measuring achievement of standards in a number of subject areas. However, we can explain a reasonable amount of the variability in NCEA results if we know an individual's cognitive (literacy, numeracy, and logical ability) and attitudinal skills (in particular the extent to which they are focused

<sup>3</sup> This calls into question the appropriateness of a linear model for a distribution that a) is severely skewed (see the histogram on the diagonal of 0), and b) has indications of heteroscedasticity. The residual plots for the model showed fewer warning signs than expected, but the results for the model should be read with some caution.

on their work and take responsibility for it). And we can explain a reasonable amount of the variability in cognitive skill by NCEA results alone.

NCEA results do allow us to distinguish between students' abilities, even if we only use very rough measures such as the number of credits they achieve or the percentage of their achievement standards that are Merit or better.