MINISTRY OF EDUCATION
Te Tāhuhu o te Mātauranga
New Zealand

Competent Learners @ 16: Competency levels and development over time — Technical report

Report to the Ministry of Education

E. Hodgen

RESEARCH DIVISION
Wāhanga Mahi Rangahau

ISBN 978-0-478-13640-1

Web Copy ISBN 978-0-478-13641-8

RMR-836

Research reports are also available on the Ministry's website: www.minedu.govt.nz/goto/2107 and on Education Counts: www.educationcounts.edcentre.govt.nz/research/index.html

Competent Learners @ 16: Competency levels and development over time — Technical report

Edith Hodgen

NEW ZEALAND COUNCIL FOR EDUCATIONAL RESEARCH

TE RŪNANGA O AOTEAROA MŌ TE RANGAHAU I TE MĀTAURANGA

WELLINGTON

JANUARY 2007

New Zealand Council for Educational Research
P O Box 3237
Wellington
New Zealand

Acknowledgements

We are very grateful for the continued involvement of the study's participants, their parents, and teachers, and their interest in the understanding that their contribution has enabled us to continue to deepen by following them through their education.

The Ministry of Education's continued funding and support has also been vital to the continuation of this project.

The fieldwork undertaken when the study participants were aged 16 took place in 2005, with most of the work having to be done in an intense effort over a six-month period. The fieldwork co-ordinators who made this possible were Cathy Lythe and Tineke Fijn, working with a great team of fieldworkers: Marion Bayne, Clare Falkner, Betty Irons, Lorraine McKay, Averill Manning, Patricia Meagher-Lundberg, Chloe Parton, Marilyn Weir, Anna Wildey, Brigid Wikinson, and Kath Wood. The data entry was undertaken scrupulously by Denise Falloon, and Christine Williams similarly did an excellent job in the final formatting of the first two reports from the age-16 phase of the project.

The project advisory group has been very helpful in our instrument design. We are particularly grateful to Sharon Cox, Heleen Visser, and Lynne Whitney from the Research Division of the Ministry of Education, and to Robyn Baker who provides NZCER's internal review, and to Sandie Aikin, Jude Allison, Dick Harker, Clive McGee, and Anne Meade.

The NZCER research team working on the age-16 phase comprises Cathy Wylie, Edith Hodgen, Rose Hipkins, Karen Vaughan, and Sally Boyd.

Table of Contents

Tables

Figures

1. Introduction

The **Competent Children, Competent Learners** study is a longitudinal study of a sample of New Zealand young people, who have been followed from their final early childhood education centres in the Wellington region. The main aim of the project is to chart the development of competencies in the context of home, leisure, and educational experiences that may account for differences in patterns of development and young people's performance. Reports from the study and associated papers are available on the NZCR website, www.nzcer.org.nz. The project is funded by the Ministry of Education, with some additional funding from NZCER.

At age 16, we have data for 448 young people. These data include results from literacy, numeracy, and logical problem-solving (pattern completion tasks), and ratings from subject teachers (English, their favourite subject, and their least favourite subject) in relation to attitudinal competencies for the 412 young people who were still at school. The latter comprise four competencies: *focused & responsible*; *thinking & learning; social skill;* and *social difficulties*. Some of these overlap with the key competencies that are now being included in the curriculum.

In this first report of the results and analysis of the material gathered during 2005, when the sample was aged 16 years, we start by describing their competency levels, and concurrent relationships between the competencies. We then turn to analyse the relationship between the young people's current competency levels, and four social characteristics: gender; family income levels; maternal qualification levels; and ethnicity. To do so, we have first compared the proportion of young people scoring above and below the median in each competency for each of the social characteristics in turn. We then fitted linear models including all of the social characteristics to see which of these characteristics contributed significantly to performance over and above the other characteristics.

Finally, we report our analysis of the predictability of current levels of performance in relation to earlier levels of performance. We have analysed the data at a number of different levels: overall trends have been modelled using structural equation models, and quartile groupings have been used to describe differences in patterns over time for high and low performers.

The **Competent Children, Competent Learners** sample was originally chosen in relation to the main focus of the first phase of the study, which was the role of early childhood education experiences and quality. This meant our units for sampling were early childhood education types, other than ngä köhanga reo, rather than social characteristics. This and the fact that our sample was chosen from the Wellington region, has resulted in a sample that is not nationally representative in terms of social characteristics. Our sample has higher proportions of young people from high-income families, and those whose mothers have trade or tertiary level qualifications, than the national average, and lower proportions of Mäori and Pacific young people.

The table below describes the sample at age 16 in terms of the four social characteristics we analyse in the study.

Table 1 ***Social characteristics of Competent Learners study sample at age 16***

	(n = 448)	***%***
Family income (at age 16)		
Low income (< $40,000)	65	15
Medium income ($40–$70,000)	122	28
High income ($70–$100,000)	90	20
Very high income ($100,000+)	143	32
Not known	23	5
Maternal qualification		
None	58	13
Trade/Mid-secondary	223	50
Senior secondary/Tertiary	80	18
University	84	19
Not known	3	1
Gender		
Male	230	51
Female	218	49
Ethnicity		
Päkehä/NZ European	360	80
Mäori	45	10
Pacific	18	4
Asian	13	3
Other	12	3

A comparison of the young people still in the study, and those who have left the study is made in Chapter 7. Up to the age of 14 (when family movements out of New Zealand were one of the main reasons for leaving the study) we found no significant differences in competency or social characteristics between those who left the study and those who were still in the study. At age 14, we found slightly lower scores for mathematics and logical problem-solving at age 12 in the group that left the study between the ages of 12 and 14. At age 16, we found more differences between those who left the study after age 14 and those still in the study at 16, and some interesting patterns over time. We also report some differences in patterns of competency levels for those who had left school by age 16, and those still at a mainstream school.

2. Age-16 competency levels for the Competent Learners sample

In this section, we look at the scores on each of our competency measures, with some description of trends in particular items making up the competency measures. The selection and development of the measures in this study used up to age 14 have been described elsewhere (Wylie, Thompson, & Kerslake Hendricks, 1996b; Wylie, 2003). Because of the period of growth covered in the young people's lives, it has not been possible to use the same measures in every phase of the study.

In this phase, we switched to a new test for reading and mathematics, and we revisited our attitudinal competency items in the light of the key competency work being done as part of the review of the national curriculum.

Cognitive competencies

We have been able to use the same test for logical problem-solving (Standard Progressive Matrices) between ages 8–16, and between ages 8–14 we used the age-related PAT standardised tests for reading comprehension, and cut-down versions of the age-related standardised PAT mathematics test. For writing, we used much the same task between ages 8 and 14, asking for a greater length at each age. We used the Burt word reading test between ages 6 and 12, but because of the ceiling effect apparent at age 12, did not use it at age 14.

At age 16 we did not use a writing test, and instead of the PAT-based mathematics and reading comprehension tests (there are no age-16 PAT tests), used a subset of the items from the International Adult Literacy Survey (IALS) developed by Statistics Canada in 1994. These tests have been used to measure adult literacy in the OECD, including New Zealand. Statistics Canada selected a subset of items that had been used previously in New Zealand, focusing on the results for 16-year-olds. The items were selected to give good discrimination between students of differing ability at age 16. Two tests were compiled, one for reading, and one for numeracy. Each test was half an hour long.

Statistics Canada converted the raw scores to a Rasch scale. They provided separate scales for prose literacy and document literacy, a combined literacy scale, and a numeracy scale. We have used the combined literacy scale and the numeracy scale in our analyses.

Mathematics/numeracy

At age 16, the IALS numeracy test consisted of 25 items based on 15 situations or pieces of information. The scale scores that Statistics Canada provided were between 114.5 and 377.7, with mean of 266.0 (s.d. 49.6). For this round of data analysis we have decided to scale all scores to be between 1 and 10 (equivalent to a tenth of a percentage), so that the parameter estimates in the models in this and later reports are easier to interpret (or at least can be interpreted more consistently). The 1–10 scaled version of the scores lie between 1.4 and 9.7, with a mean 6.0 (s.d. 1.5).

How do the PAT-based tests (particularly that administered at age 14) compare with the IALS test?[1]

- Both tests are appropriately aimed at Level 5 of the New Zealand mathematics curriculum, making them suitable for the large majority of students in the study who had completed Year 10
- The general level of difficulty is about the same, but the IALS test has items requiring more sophisticated statistical interpretations, and probabilistic and proportional reasoning.
- The length of the questions (the number of words used) is about the same, but some of the IALS questions have longer stems
- The IALS test is more grounded in real context and solutions to everyday problems, being more a test of numeracy than the PAT test which tests more mathematical knowledge and skills
- The IALS test has constructed responses; the PAT test is multiple choice. This means that the item difficulty of the IALS test will be slightly higher (the students are free to give more wrong answers, and have no guides as to what the correct answer might be), and the test may take longer to complete
- The IALS test has a higher language load, both in the longer, more wordy items, and in the written explanations required for some questions. Some questions require extensive additional information to be absorbed and extracted from the text
- The IALS test has more measurement/geometry questions and statistics questions; the PAT test has more number (fractions/proportions, pattern-finding, pre-algebraic reasoning) questions
- The IALS test has a common context and stem for several questions, which reduces the independence between items, and may have an effect on the reliability of the test.

Given the overall amount of similarities and differences, we would expect similar results from the two tests and that the results would be highly correlated. But the differences are enough to think that the correlation between the age-14 and age-16 scores may be slightly lower than that between the age-12 and age-14 scores.

Literacy

The IALS literacy test consisted of 27 items based on nine pieces of prose of varying length and complexity. The scale scores that Statistics Canada provided were between 164.9 and 357.9, with mean 280.7 (s.d. 38.73). The 1–10 scaled version of the scores lie between 2.9 and 9.7, with mean 6.4 (s.d. 1.2).

What are the similarities and differences between the PAT reading comprehension tests and the IALS test?

- The IALS test requires written answers, while the PAT test is multiple choice
- Most IALS test tasks are information retrieval
- The IALS texts have a variety of features not found in the PAT tests, such as bullet pointed information, structured text with headings and sub-headings, extracts from brochures
- Some of the IALS questions are markedly harder (longer and more complex text, more challenging questions) than others, while the PAT questions tend to be more similar and more consistent across the test

[1] My thanks to the NZCER colleagues who used their curriculum and test development knowledge to compare the tests: Charles Darr and Alex Neill (mathematics), and Juliet Twist and Verena Watson (reading).

- This meant that to obtain a high score on the IALS test it was important to be able to skim read, it was advantageous to read the question first and then search for corresponding key words in the text, particularly as the answers to successive items set on a particular portion of text were not necessarily to be found in the same order in the text (answers to the first three items may be found in paragraphs 7, 4, and 11, say)
- Respondents to the IALS test are seldom asked to analyse, and may find the occasional change, from information seeking to analysis, challenging
- PAT tests, being multiple response, give clear indications to the respondent how many pieces of evidence[2] are expected for each item; some of the items in the IALS test did not give clear indications as to how much detail was expected.

As for the mathematics/numeracy scores, we would expect there to be strong correlations with the age-14 score, but possibly not quite such a strong correlation as between the age-12 and age-14 scores.

Logical problem-solving

The average raw score for the sample was 48.3 (s.d. 6.7). The lowest score was 17 items of the 60 test items correct, and the highest, 59 items.

Comparison of the results across ages 12, 14, and 16 indicated that some individuals had widely varying scores in the three data collection rounds, and in the age-16 round there were individuals who had very low logical problem-solving scores, but relatively high literacy and/or numeracy scores. This is a pattern that has not been observed in the past, and at this round was sufficiently strong to raise the question as to whether these individuals were "outliers" and would affect later analyses.

Our approach at this age started by looking at patterns over the last three phases of the study. The test is such that the scores increased fairly rapidly with age in the early years, but this rate of increase has slowed since age 12, so that we would expect the scores to converge after that age. Most of the young people did have similar (convergent) scores, but a few did not. To attempt to smooth this effect over time, giving a single, consistent and representative measure for this competency, we calculated weighted mean scores. So that the scores reflected the recent past more strongly than the more distant past, a weighting of 1 : 2 : 2.5 was used for the age-12, -14, and -16 scores, respectively. The mean score has increased as the young people grew older, and to compensate for this each raw score was standardised (to mean 0 and standard deviation of 1), the weighted mean was calculated, and the resultant score was then scaled back to be on a 1–10 scale. The 1–10 scaled version of the scores lie between 3.6 and 9.9, with mean 8.0 (s.d. 1.0).

Attitudinal competencies

The teachers of those still at school at age 16[3] were asked to characterise the study students. Each student was asked to nominate their favourite and least favourite subjects (or second best/worst if either was English), and the

---

[2] For example: "What factors did …" as opposed to "What three factors did …"

[3] Twenty-seven of the participants at age 16 had left school, and five were in non-mainstream schools. Where possible, teacher ratings of these five were used for the attitudinal competencies, but there was more missing data than for other students.

English, favourite subject, and least favourite subject teachers were asked to complete the questionnaire. We had a total of 1249 responses (plus one that was blank as the teacher was at camp), from 418 English teachers, 417 favourite subject teachers, and 414 least favourite subject teachers.

The broad groupings of subjects nominated as favourite and least favourite are shown in Table 2. It is noticeable that there is a greater spread of opinion about which subject the students enjoyed most than there is about which subject they enjoyed least. A quarter nominated a science subject, and almost 40 percent nominated mathematics as their least favourite subjects. These are the most "academic" of the subjects.

Table 2 ***Most and least favourite subjects at age 16***

Subject	***Nominated as favourite (n = 415) %***	***Nominated as least favourite (n = 417) %***
Arts	19	2
Health/Physical education/Food technology	17	4
Sciences	13	25
Technology	10	4
Humanities	10	11
Applied work/study skills	7	3
Mathematics	6	37
Graphics	4	1
Computer studies/Information technology	4	4
Business studies	3	6
Languages	3	3
Other	2	1

The responses from the English teachers were the most complete; teachers who had less contact with the student (physical education, in some schools, or where the student was on the roll but not attending class), or who had more formal and structured classes (mathematics and science subjects), or who did not have face-to-face contact with the students (those at The Correspondence School, about four teachers) tended to find some of the questions hard to answer. The items that had the highest rates of non-response (10–25 percent of respondents) were those that required insight into the character of the student, or knowledge of their behaviour (in or outside the classroom). The items with the lowest rates of non-response were those that were simple and factual: whether the student asked for advice or clarification, finished their work, had good concentration, or arrived on time.

Not all of the items used to construct the competencies at age 14 were asked again at age 16, and some new items were included in this round, in relation to the new key competencies in the revised curriculum. This meant that using the same competencies as age 14 was impossible, so we developed new competencies.

An examination of the data indicated that there were marked differences between the responses by the three teachers: the responses by the favourite subject teacher were more favourable; those by the least favourite subject teacher were less favourable; and those by the English teacher were between the others. This was true for most of

the items. A preliminary factor analysis[4] of all the items for all of the teachers together indicated that the responses by the separate teachers would load on different factors, so to investigate the possible scale or factor structure three separate factor analyses were run for the three different sets of teachers. There was good agreement between the three that four factors could be formed, and the items loading on each factor were largely consistent.

However, if we used separate factors for each teacher, we would have 12 competencies, each with between 8 and 33 "missing values" in the sense that there was insufficient information to obtain a scale score for between 8 and 13 students. If the competency was used in later analysis, this is how many of the students would be excluded from the analysis. Combining the items across teachers looked more attractive, and had the advantage of smoothing out the more and less positive views of the most and least favourite teachers. We found that combining the factor items across teachers gave four new competency scales that worked relatively well (details are given below), and that had eight or nine missing values, and which related well to earlier competency measures.

To form the competencies, we calculated the (unweighted) mean across the items for all three teachers so long as we had responses to at least half of the items concerned. The reliability of these scores, as measured by Cronbach's alpha, lay between 0.79 and 0.97.

Thinking & learning

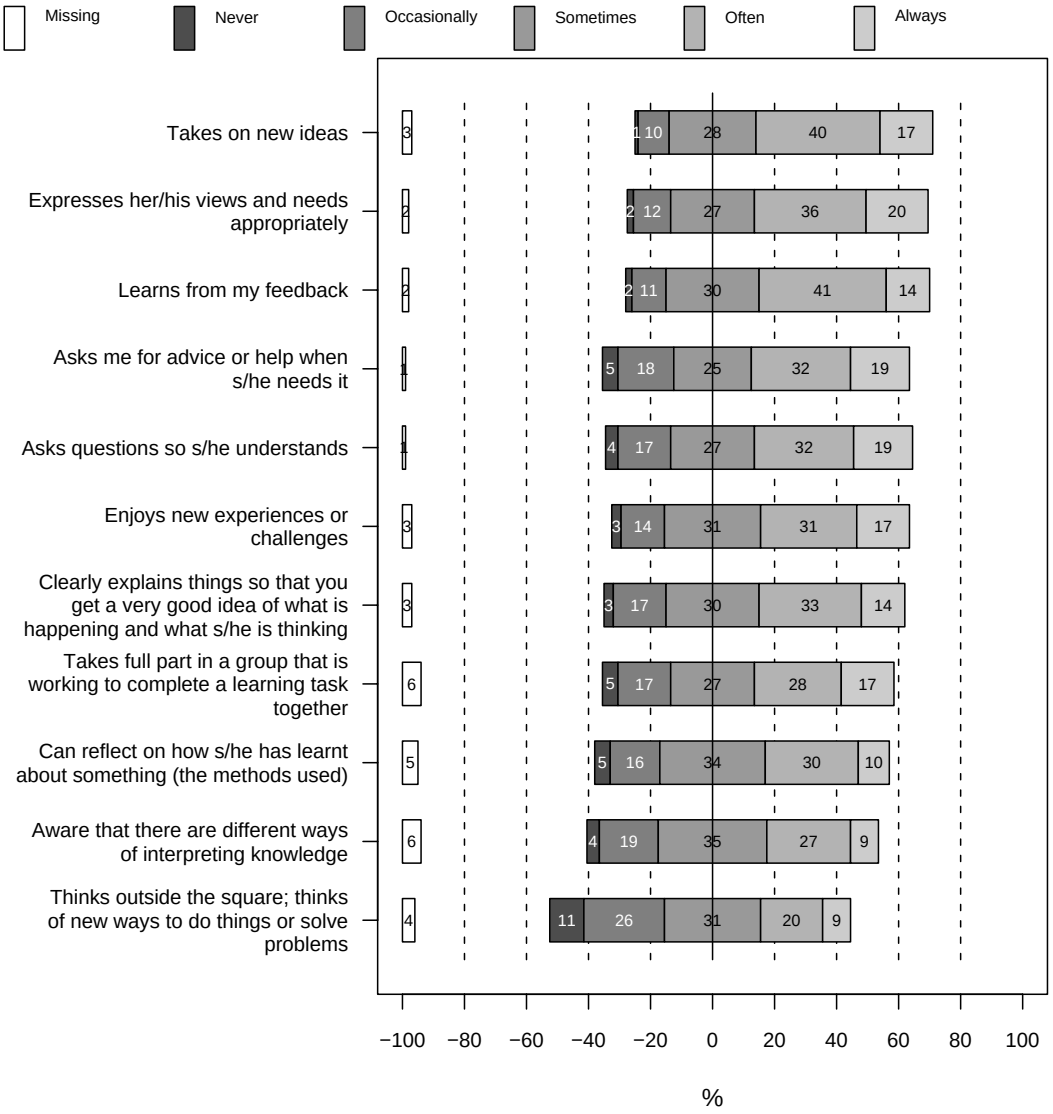
This scale uses items from the age-14 *curiosity* and *communication* scales as well as some of the new items, and had a reliability measure of 0.96 (Cronbach's alpha). The correlation with the age-14 *communication* scale is 0.59, and with age-14 *curiosity* is 0.56, both considerably stronger than the corresponding correlations between the age-14 and age-12 competencies (0.44 and 0.40, respectively).

The new scale, from a possible 36 items (12 items from each of three teachers), was constructed for all students for whom we had responses to at least 18 of the items. The correlations between the individual items and the scale score ranged between 0.51 and 0.68, and 21 of them were 0.60 or more. The scores on the 1–10 scale were between 1.6 and 10, with mean 6.3 (s.d. 1.5).

The topmost three bars in Figure 1 show that around half of the age-16 students in our sample were seen by their teachers to often or always show openness to new ideas (57 percent) and feedback (55 percent), and to be able to articulate their views and needs (56 percent). This sample of students was somewhat less likely to be reflective in their learning (40 percent), or think outside the square (29 percent). The proportion of those who did think outside the square often or always was much the same as it had been when the students were 14.

[4] Principal factor analysis with varimax rotation, using SAS/STAT® software, Version 9 of the SAS System for Windows.

Figure 1 ***Responses to items on the thinking & learning scale***



Focused & responsible

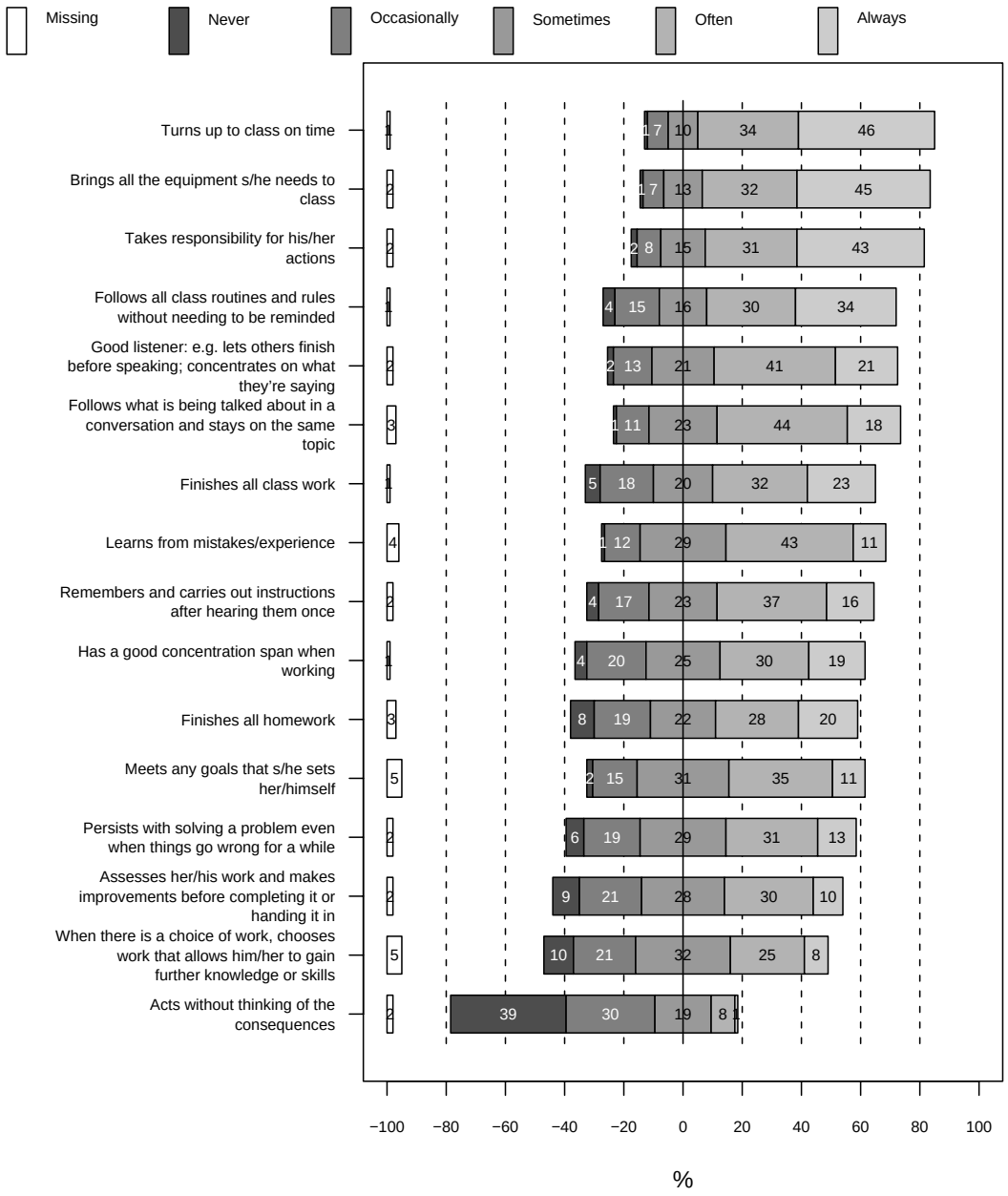
This scale uses items from the age-14 *perseverance*, *self-management*, *self-efficacy*, and *communication* scales, and had a reliability measure of 0.97. The correlations with the age-14 scales are 0.69, 0.67, 0.53, and 0.61 respectively. These compare well with the age-12 and age-14 correlations of 0.56 for *perseverance*, and 0.44 for *communication*.

The new scale was constructed from a possible 48 items (16 items from three teachers), and was calculated for all students for whom we had responses to at least 24 of the items. The correlations between the individual items and the scale score ranged between 0.50 and 0.75, with five of them under 0.60, and nine that were 0.70 or more. The scores on the 1–10 scale were between 2.8 and 9.8, with mean 6.8 (s.d. 1.6).

On the whole, as is shown in Figure 2, the teachers' views show that the age-16 students were reasonably well organised. More than three-quarters of the age-16 students often or always turned up to class on time (80 percent), brought all the equipment they needed (77 percent), and took responsibility for their own actions (74 percent). They did not do everything asked of them: just over half often or always finished all their class work (55 percent);

and just under half, their homework (48 percent). In terms of stretching or challenging themselves, 40 percent assessed their own work and made improvements to it before handing it in, and around a third chose work that allowed them to gain further knowledge or skills.

Figure 2 ***Responses to items on the focused & responsible scale***



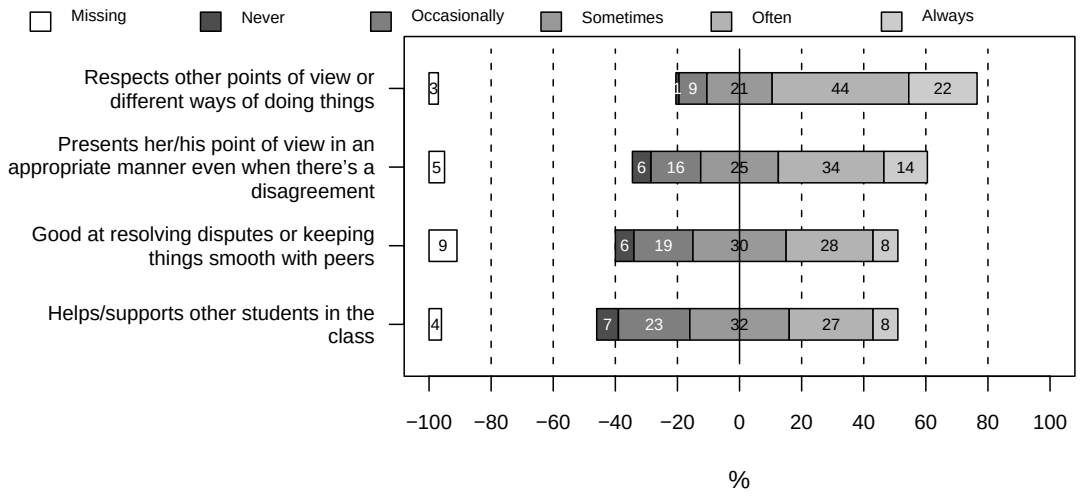
Social skills

This scale uses some items from the age-14 *social skills with peers* scale, and had a reliability measure of 0.79. The correlation with the age-14 scale is 0.47. This compares well with the age-12 and age-14 correlation of 0.38.

The new scale was constructed from a possible 12 items (four items from three teachers), and was calculated for all students for whom we had responses to at least six of the items. The correlations between the individual items and the scale score ranged between 0.37 and 0.51, with two of them under 0.40, and three that were 0.50 or more. The scores on the 1–10 scale were between 1.6 and 10, with mean 6.3 (s.d. 1.6).

Two-thirds of the age-16 students were seen to respect differences in others views or ways of doing things (Figure 3). Some had more active social skills: around a third often or always were good at resolving disputes among peers, or helping other students.

Figure 3 ***Responses to items on the social skills scale***



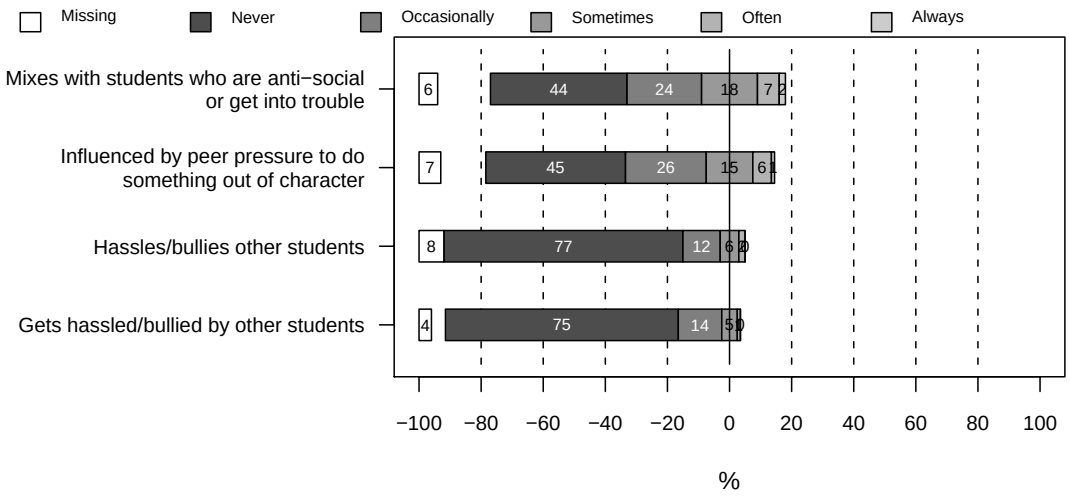
Social difficulties

This scale identifies students who are involved in bullying, and/or have friends who may be a bad influence. On the other attitudinal scales a high score is "good", and the items for this scale have been reversed, so that the score for this competency, too, conforms to the "high is good" rule. A *low* score identifies those with difficulties.

This scale uses some items from the age-14 *social skills with peers* scale, and had a reliability measure of 0.79. The correlation with the age-14 *social skills with peers* scale is 0.50. This compares well with the age-12 and age-14 correlation of 0.39.

The new scale was constructed from a possible 12 items (four items from three teachers), and was calculated for all students for whom we had responses to at least six of the items. The correlations between the individual items and the scale score ranged between 0.17 and 0.57, with two of them under 0.30, and four that were 0.50 or more. The scores on the 1–10 scale were between 4.6 and 10, with mean 8.7 (s.d. 1.1). The scores on this scale are markedly skew (most students do not have a low risk profile using this scale), so for further analysis it may be more useful to identify the students most at risk.

Figure 4 ***Responses to items on the social difficulties scale***



Summary

At age 16 we have three individual and one composite measures of cognitive competency that are similar to those used previously in this study. The three separate competencies are: *numeracy*; *literacy;* and *logical problem-solving*. We have these measures for all young people still in the study who were willing and able to complete the tasks (about 444 young people).

In addition, for the 412 students still in a mainstream school we have four measures of attitudinal competency, plus a composite attitudinal score. The four separate scores were built from combinations of the items used to form the age-14 attitudinal competencies, plus some new items introduced in this round, more relevant for 16-year-olds. The competencies are: *focused & responsible; thinking & learning; social skills;* and *social difficulties*.

3. Concurrent relationships between competency measures at 16

In this section, we look at the correlations between the competency measures, then at some transformations for some of the competencies, to allow analysis, and then we fit a structural equation model to the age-16 measures.

Correlations

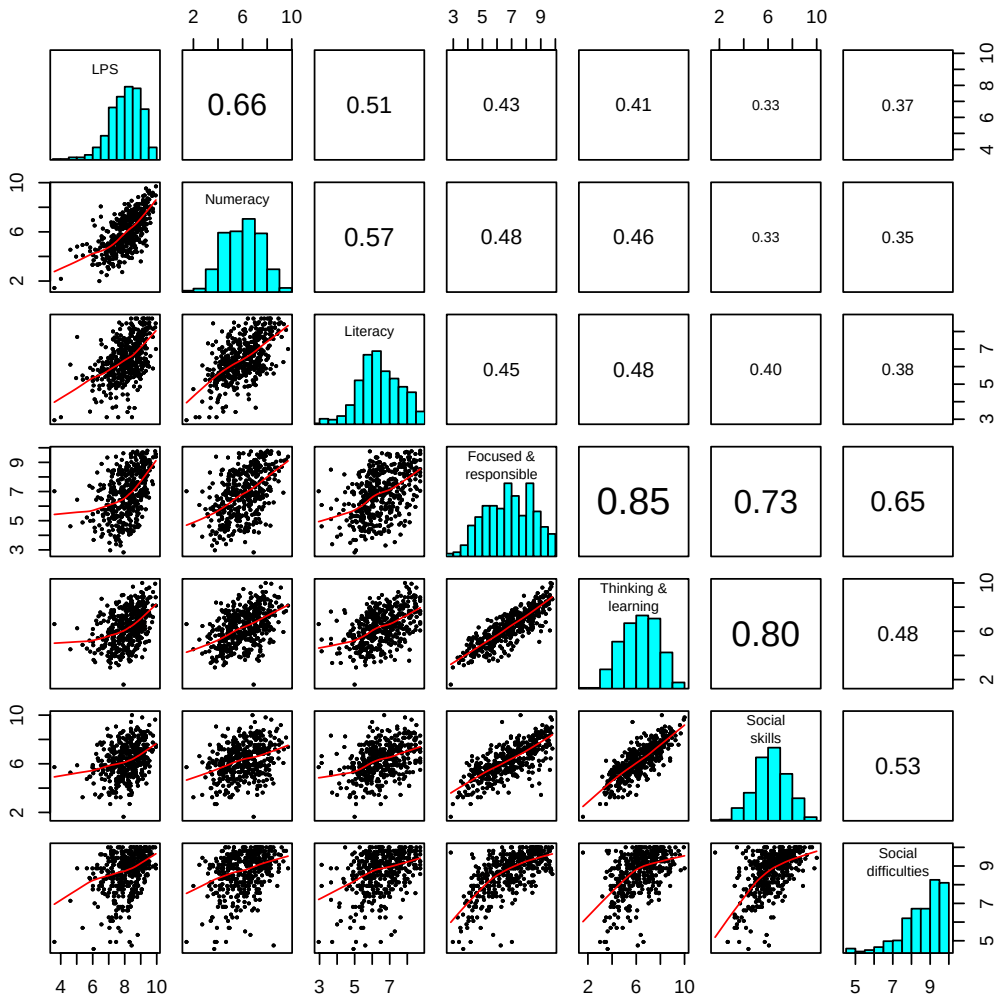
There are strong associations between the competencies at age 16. The strongest correlations are between *thinking & learning* and *focused & responsible* ($r = 0.85$), and between *thinking & learning* and *social skills* ($r = 0.80$). All the correlations, and plots of the associations are shown in Figure 5.

The numbers in the upper triangle of the plot are the Pearson's product-moment correlations, and the corresponding scatter-plots are shown in the lower triangle. For example, the correlation between *numeracy* and *focused & responsible* is 0.48 (second row from the top of the plot for *numeracy* and fourth column from the left for *focused & responsible*), and the plot corresponding to this is in the fourth row from the top for *focused & responsible* and second column from the left for *numeracy*. To see which axis on the plot belongs to which competency, look *up* the column of plots to see the competency on the horizontal axis, and *across* the row of plots to see the competency on the vertical axis. So to continue the example above, *numeracy* is on the horizontal axis and *focused & responsible* is on the vertical axis. On the diagonal of the plot, along with the name of the competency, is a plot showing the distribution of the scores for each competency. The size of the font used to represent the correlation coefficients is proportional to the size of the correlation coefficient. So stronger correlations have larger numbers, and weaker correlations have smaller numbers.

The lines through the scatter-plots give insight as to whether the relationship between the pairs of competencies is linear (in which case the line is more or less straight) or not.[5] Much of the modelling that we do is based on the assumptions of a linear (straight-line) relationship, so where there are strong indications on non-linearity it is best to transform the data so that the relationship is more nearly linear (see the next section). Some of the curved lines are affected by a few outliers, for example in *logical problem-solving* (abbreviated to LPS in the figure) and *focused & responsible*, and others represent an actual curvilinear relationship, for example *focused & responsible* and *social difficulties*.

[5] A local linear regression smoother was used to plot the lines.

Figure 5 ***Associations between competencies at age 16 (n = 448)***



Transforming the data

The scatter-plots in Figure 5 above indicate that at least some of the variables should be transformed before being used in a model in which multivariate normality is assumed. Box-Cox transformations are straightforward and commonly used. These transformations are power transformations with the added advantage that they converge to a log transformation as $a \to 0$. The transformations are of the form:

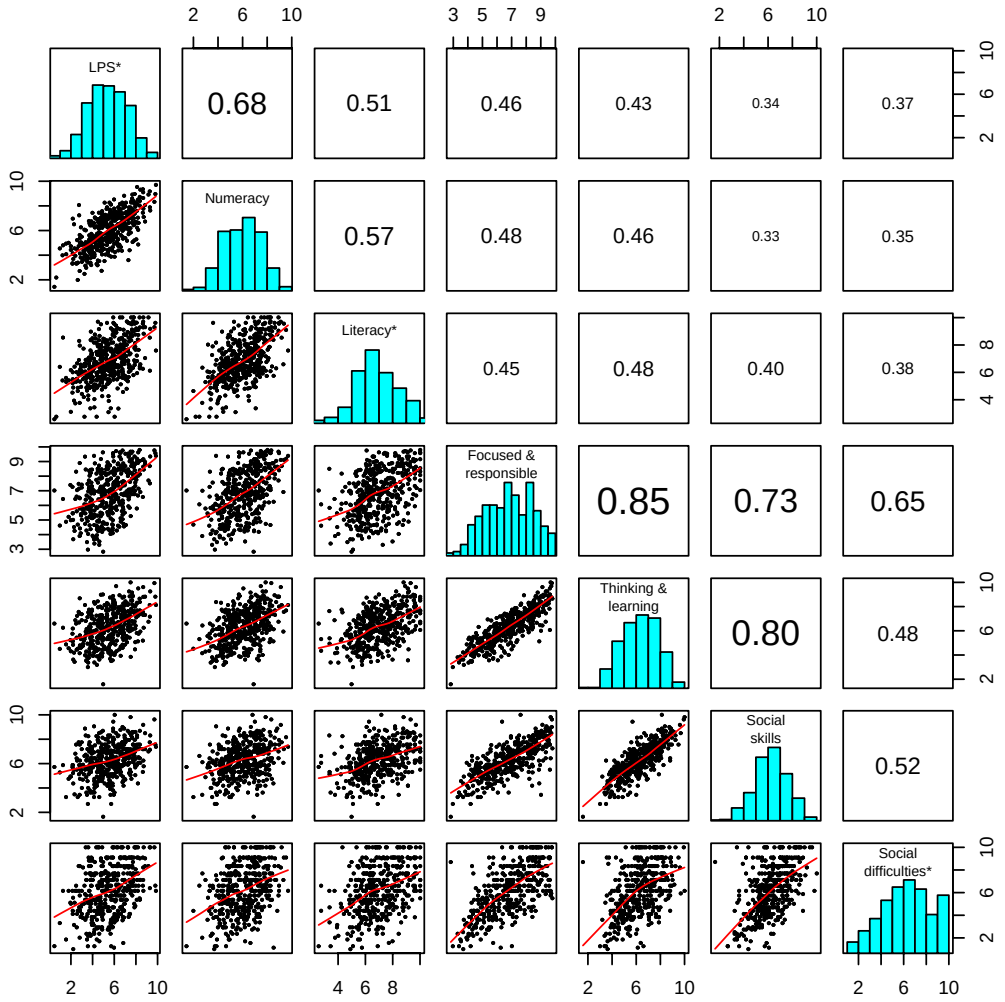
$$y^* = \begin{cases} \dfrac{y^a - 1}{a} & \text{for } a \neq 0 \\ \log(y) & \text{for } a = 0, \end{cases}$$

where the value of a chosen for each variable brings the distribution of y as close as possible to normal. The competencies that benefited from the transformation were *logical problem-solving* (the value of a used was 3), *literacy* (the value of a used was 1.25), and *social difficulties* (the value of a used[6] was –0.1). The effect of the

[6] This was the value used on the original mean of the items, in which a high score indicated greater levels of involvement in bullying or other forms of trouble.

transformations can be seen in Figure 6, where there are fewer curved lines through the scatter-plots. Increasing the linearity of the relationships has the effect of increasing some of the correlations, because correlation is a measure of straight-line association (a curvilinear relationship will always have a lower correlation coefficient than an equally strong linear relationship).

Figure 6 ***Linear associations between competencies at 16 after transformation (n = 448)***



* indicates the competencies that have been transformed.

Structural equation model of relationships between competencies at age 16

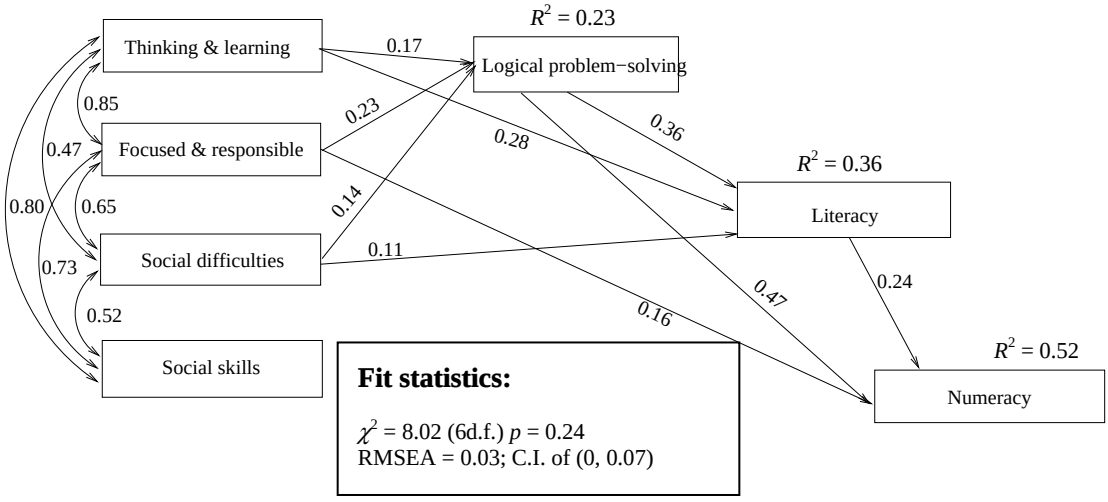
At age 14 we could build a model in which the attitudinal competencies were used to predict *logical problem-solving*, *PAT reading comprehension*, and *mathematics*, and the first two of the cognitive competencies were used to predict *mathematics*.

At age 16 we could build a remarkably similar model, shown in Figure 7.[7] Strictly speaking, this is a path model, as it does not involve any latent variables (all are "manifest" or observed, although that is not strictly true, as they are derived from responses to items; however the item responses were not used to form the model, only the scale

[7] We used PROC CALIS, SAS/STAT software to fit the model.

variables constructed from them). All the paths shown in the model make a statistically significant contribution to the model.

Figure 7 ***Structural equation model of relationships between competencies at age 16***



The attitudinal competencies are the exogenous (explanatory) variables: *thinking & learning*; *focused & responsible*; *social difficulties*; and *social skills*. These variables are all strongly correlated, as is shown by the correlation coefficients on the far left of the diagram. The strongest correlation of 0.85 is between *thinking & learning* and *focused & responsible*, and the next strongest of 0.80 is between *thinking & learning* and *social skills*. Note that these correlations provided by the model are the same as those shown in Figure 6 above, apart from slight rounding errors. Some of these correlations are so high that there was a possibility of multicollinearity in the model. However, there were no signs of the effects of multicollinearity, apart from the fact that once *thinking & learning* was used to predict *logical problem-solving* and *numeracy*, paths from *social skills* to the same cognitive competencies were not significant. Eliminating *social skills* from the model altogether had no effect on the path coefficients (did not affect the model fitted), but did affect the goodness-of-fit statistics slightly in a way that did not improve the model. Therefore the competency has been retained in the model, but its only influence is indirect, through its correlation with the other attitudinal competencies.

The paths, shown by straight arrows, together with their associated coefficients (just above or below the path arrow) indicate the extent to which the variable at the end of an arrow is explained by the variable at the beginning of the arrow (in this model). All the influences are relatively weak (for example, between *social difficulties* and *literacy*, a coefficient of 0.11) to moderate (for example, between *logical problem-solving* and *numeracy*, a coefficient of 0.47).

The R^2s shown by each of the endogenous (outcome) variables are a measure of how well the model explains those variables. A number close to 1 would indicate that the model is explaining that variable well, and a number closer to 0 would indicate that the variable is not being explained well. The model explained the *numeracy* score better than the *logical problem-solving* score (0.52 is higher than 0.23).

The goodness-of-fit statistics quoted are the χ^2 goodness-of-fit statistic and the RMSEA (root mean square error of approximation). The χ^2 statistic compares the observed and modelled covariance matrices, and tests the null hypothesis that they are equal. In this instance the fit is good (p-value well over 0.05). The RMSEA should have a

value of less than 0.06 and an upper confidence limit of at most 0.065. The first criterion is satisfied, and the second very nearly is.

The model suggests that the *numeracy* score is accounted for directly by the *focused & responsible* score, *logical problem-solving* score (the strongest of the effects), and *literacy* score, and indirectly by the other attitudinal competency scores which have direct effects on *logical problem-solving* and/or *literacy*.

The *literacy* score is accounted for directly by the *thinking & learning social difficulties*, and *logical problem-solving* scores, and indirectly by the other attitudinal competency scores which have direct effects on *logical problem-solving*.

In the model fitted, *thinking & learning* has only an indirect effect on *numeracy*, and *focused & responsible* has only an indirect effect on *literacy*. This pattern makes some sense when we think of the different items in the two competencies. Those with a high score in *thinking & learning* were rated highly by their teachers in things relating to the use of language, and expressing themselves well. They are therefore likely to be interested in words, their meaning, and the ideas they can convey, and to score highly in any literacy test. Those with a high score in *focused & responsible* are orderly, self-disciplined, stick to tasks, and get satisfaction from and take pride in task completion. They are likely to find fascination in and be rewarded by the orderly world of numbers, and so to do well in any mathematics or numeracy test.

Of course, the correlation between *thinking & learning*, and *focused & responsible* was high (0.85), and that between *literacy* and *numeracy* was moderately high (0.57). The correlation between *focused & responsible* and *numeracy* (0.48) was only slightly higher than that between *focused & responsible* and *literacy* (0.45), and those between *thinking & learning* and each of *numeracy* and *literacy* had similar values (0.46 and 0.48, respectively). So there is not much in the pattern of paths in the diagram, but the pattern, such as it is, does make sense.

Path models like this one can be used to demonstrate that possible causal path models fit the observed data well. However, causality and a good fit do not always go hand-in-hand. In the present model, the way the variables are fitted suggests that social difficulties "caused" low cognitive competency results. However, it may well be that students who have lower levels of logical problem-solving and lower literacy levels, are more likely to become involved in behaviours that were counted as contributing to "social difficulties", rather than the other way around.[8] Bear in mind, too, that this model is focused on relationships between current competency levels; we turn to the relations of previous competency levels with current competency levels in Chapter 5. But before then, we look at whether students from different social backgrounds show different patterns of performance at age 16.

Summary

There are moderate to strong associations between the cognitive and attitudinal competencies, with the strongest associations being between the attitudinal competencies. So far as the cognitive competencies are concerned, the young people with a strength in one area were likely to have strengths in other areas, although there were individuals with a particular strength in one or two areas, but not all of them. However, for the attitudinal

[8] An attempt to fit a different formulation of the model, where *social difficulties* became one of the endogenous (outcome) variables appeared to give a better fit, but also gave stronger indications of the effect of multicollinearity and so was not pursued further.

competencies, particularly *focused & responsible*, *thinking & learning*, and *social skills*, most young people had approximately equal strengths in all areas.

We were able to fit a path model in which the correlated attitudinal competencies explained 23 percent of the *logical problem-solving* score (in the model the influence of *focused & responsible*, *thinking & learning*, and *social difficulties* was direct, that of *social skills* was indirect); the *logical problem-solving*, *thinking & learning*, and *social difficulties* scores explained 36 percent of the *literacy* score; and the *logical problem-solving*, *focused & responsible*, and *literacy* scores explained 52 percent of the *numeracy* score.

4. Differences in age-16 performance related to social characteristics

Over the successive rounds of data collection we have traced possible differences in achievement between groups of young people. We have used demographic variables to define groups, not because we think that the young people may perform differently *per se* because of the groups to which they belong, but because we think that those within the groups may have had common circumstances or experiences that may be associated with differing levels of achievement. The demographic variables used for this round of data collection are: maternal qualification,[9] age-5 family income,[10] gender, and ethnicity (of the young person, as recorded at age 5).[11]

An initial, simple way to compare groups is to look at the percentage in each group scoring above the median for each competency. We have cognitive competency scores for both the young people still at school, and those who have left school, and attitudinal competency scores for those still at school only, as these scores are based on teachers' evaluations of the young people.

Above median competency scores and social characteristics

Table 3 shows clear linear trends in relation to the cognitive competencies for maternal qualification and family income levels, and similar but weaker trends for ethnicity. Gender differences are statistically significant for *literacy*, but not for *numeracy* or *logical problem-solving*.

---

[9] Note that between writing the equivalent report on competencies at age 14 and the rest of the reports at age 14 we redefined the categories of maternal qualification. In this report we are using the categories used in the bulk of the age-14 reports. The first and last categories (no qualifications and university qualifications) are the same in the two categorisations used, but whereas up to age 12 we had a category for all those with school qualifications, and one for all those with trade or tertiary qualifications, the later age-14 reports and age-16 reports have a category for those with mid-school or trade qualifications, and one for those who have senior secondary school or tertiary qualifications. We recategorised because of the similarities in competency levels between the different groups.

[10] We use family income levels at the early stage of the young people's lives because early family income has often shown more of a contribution to competency levels than current family income, and to keep consistency with previous analyses.

[11] For consistency with earlier reports, we here use age-5 ethnicity. At age 16 we asked the young people to say which ethnic group(s) they identified with, and how the groups specified at age 5 and those given at age 16 differ will be discussed elsewhere.

Table 3 ***Percentage of age-16 young people scoring above the median in the cognitive competencies***

Competency measure → ***Group ↓***		***Logical problem-solving (n = 446)***	***Numeracy (n = 444)***	***Literacy (n = 444)***
Maternal qualification				
None	(*n* = 57)	26	23	28
Mid-school/Trade	(*n* = 220)	42	46	47
Senior secondary/Tertiary	(*n* = 80)	65	55	54
University	(*n* = 83)	75	77	71
p-value for χ^2 tests		< 0.0001	< 0.0001	< 0.0001
Family income group at age 5				
Under $30,000	(*n* = 108)	36	31	32
$30,000 to $60,000	(*n* = 198)	49	51	51
$60,000 to $80,000	(*n* = 61)	62	64	66
Over $80,000	(*n* = 64)	64	67	64
p-value for χ^2 tests		0.0019	< 0.0001	< 0.0001
Gender				
Female	(*n* = 217)	51	46	59
Male	(*n* = 227)	49	54	42
p-value for χ^2 tests		0.778	0.087	0.0006
Ethnicity				
Päkehä/Asian[a]	(*n* = 369)	52	53	52
Mäori/Pacific[b]	(*n* = 63)	37	29	35
p-value for χ^2 tests		0.044	0.0011	0.0216

[a] 360 Päkehä and 13 Asian young people.

[b] 45 Mäori and 18 Pacific young people. The two groups shown here bring together the ethnic categories whose age-14 competency levels were most similar, so that we had groups of sufficient size for comparison. The numbers of Mäori and Pacific young people left in the study at age 16 are lower than desirable. However, our caveats about this low number are tempered by the fact that our findings for this sample are consistent with other studies of Mäori and Pacific students' performance.

The patterns found at age 16 are strongly similar to those found at age 14. The largest, most marked differences are those for the competencies learnt at school (*numeracy* and *literacy*), across groups defined by "advantage" (maternal qualifications and family income). This "advantage" would include patterns of interaction and experience (e.g. having more experiences of books and activities at home that support the development of literacy and numeracy), and the consequences of socioeconomic advantage such as having access to higher-decile schools, offering peers from similarly advantaged homes.

Similar trends were found for the attitudinal competencies (Table 4). There were statistically significant differences between all groups for all competencies, except for *social skills* and *social difficulties* for the ethnic groups, and *social difficulties* for the family income groups.

Table 4 ***Percentage of age-16 young people scoring above the median in the attitudinal competencies***

Competency measure → ***Group ↓***		***Thinking & learning (n = 412)***	***Focused & responsible (n = 412)***	***Social skills (n = 412)***	***Social difficulties[a] (n = 412)***
Maternal qualification					
None	(*n* = 48)	27	29	42	38
Mid-school/Trade	(*n* = 205)	45	42	43	48
Senior secondary/Tertiary	(*n* = 76)	57	59	61	62
University	(*n* = 80)	70	73	64	79
p-value for χ^2 tests		< 0.0001	< 0.0001	0.0013	< 0.0001
Family income group at age 5					
Under $30,000	(*n* = 87)	39	32	41	45
$30,000 to $60,000	(*n* = 191)	48	52	49	56
$60,000 to $80,000	(*n* = 59)	66	59	70	58
Over $80,000	(*n* = 64)	56	56	50	67
p-value for χ^2 tests		0.0150	0.0057	0.0075	0.0847
Gender					
Female	(*n* = 198)	57	56	59	66
Male	(*n* = 214)	43	44	42	46
p-value for χ^2 tests		0.0058	0.0140	0.0008	< 0.0001
Ethnicity					
Päkehä/Asian	(*n* = 344)	54	54	52	57
Mäori/Pacific	(*n* = 56)	29	27	41	43
p-value for χ^2 tests		0.0011	0.0009	0.3012	0.0970

[a] A score above the median in this competency is indicative of *low* risk responses (is "good").

Variation in competency scores related to social characteristics

The tables above give a broad-brush picture of social group differences in competency levels. We get a more detailed look at the information if we illustrate the differences using box plots (we see the actual distribution of the scores for each group).

The features of the box plots used are:

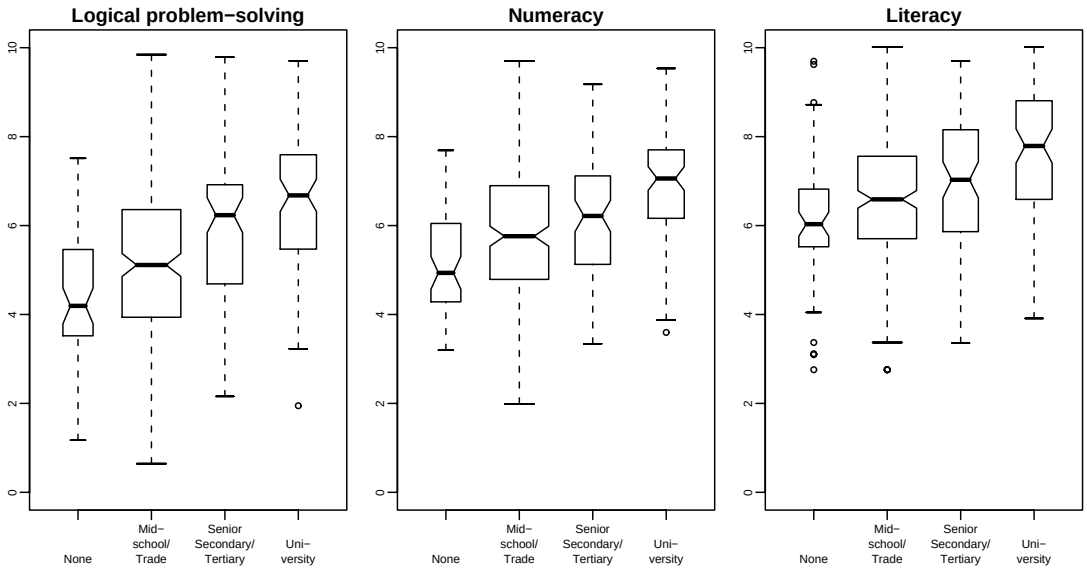
- The width of each box is proportional to the number of young people in that group (the sample size)
- The vertical axis gives the score achieved for the test
- The bars at the top and bottom of the plot show the values of the highest and lowest scores respectively
- Where there are extremely high (or low) scores, these outliers are shown by circles above (or below) the maximum (or minimum) score bars

- The dashed whiskers show the range of scores achieved by the top 25 percent of the young people and the lowest 25 percent of the young people (the first and fourth quartile groups, respectively)
- The box shows the scores achieved by the middle 50 percent of the young people (the second and third quartile groups, respectively), with horizontal lines at the top and bottom of the box at the first and third quartiles
- The narrower but fatter line at the centre of the box shows the value of the median. If the distribution of the scores is symmetric, the median is at the centre of the box and the whiskers are more or less the same length. If the distribution of the scores is skew, then the median is closer to one quartile than to the other, and, typically, the whisker on the "short quartile" side is shorter and the other whisker is longer
- The notch on the side of the bars indicates an approximate 95 percent confidence interval for the median. If the notched sections on two boxes do not overlap, then it is probable that the competency scores for the two groups defined by those boxes are significantly different. This is only a rough guide, but can give a fairly good indication of where there are (or are not) differences.

In the discussion of each of the sets of box plots that follow, mention is made of groups that are "likely" to be statistically significantly different for each competency. This relatively informal approach is supplemented in the next section, where a model is fitted for each competency with all the social characteristics as explanatory variables. In the tables in that section the levels of each social characteristic that did differ significantly for the competency are listed.

The scores used to make the box plots were the transformed scores (see the previous chapter) where appropriate: in *logical problem-solving*, *literacy*, and *social difficulties*. The cognitive scores include those of the students who have left school, and the attitudinal scores are those of the students still at a mainstream school only.

Figure 8 ***Cognitive competencies measured on 1–10 scales for maternal qualification groups***



We can see from Figure 8 that the variation in cognitive competency between groups has the characteristics:

- The highest and lowest scores in each group are relatively similar (although the highest scores for the group with mothers with no formal qualifications were relatively lower in *logical problem-solving* and *numeracy*)

- There is, however, a relatively strong gradient *on average* across the groups, with the median scores increasing with increasing maternal qualification
- The greatest "jumps", probably indicative of statistically significant differences, are between the no qualifications group and the mid-school or trade qualifications group, and between the secondary school or tertiary qualifications group and the university qualifications group. This is most marked for the *literacy* and *numeracy* scores.

Figure 9 ***Attitudinal competencies measured on 1–10 scales for maternal qualification groups***

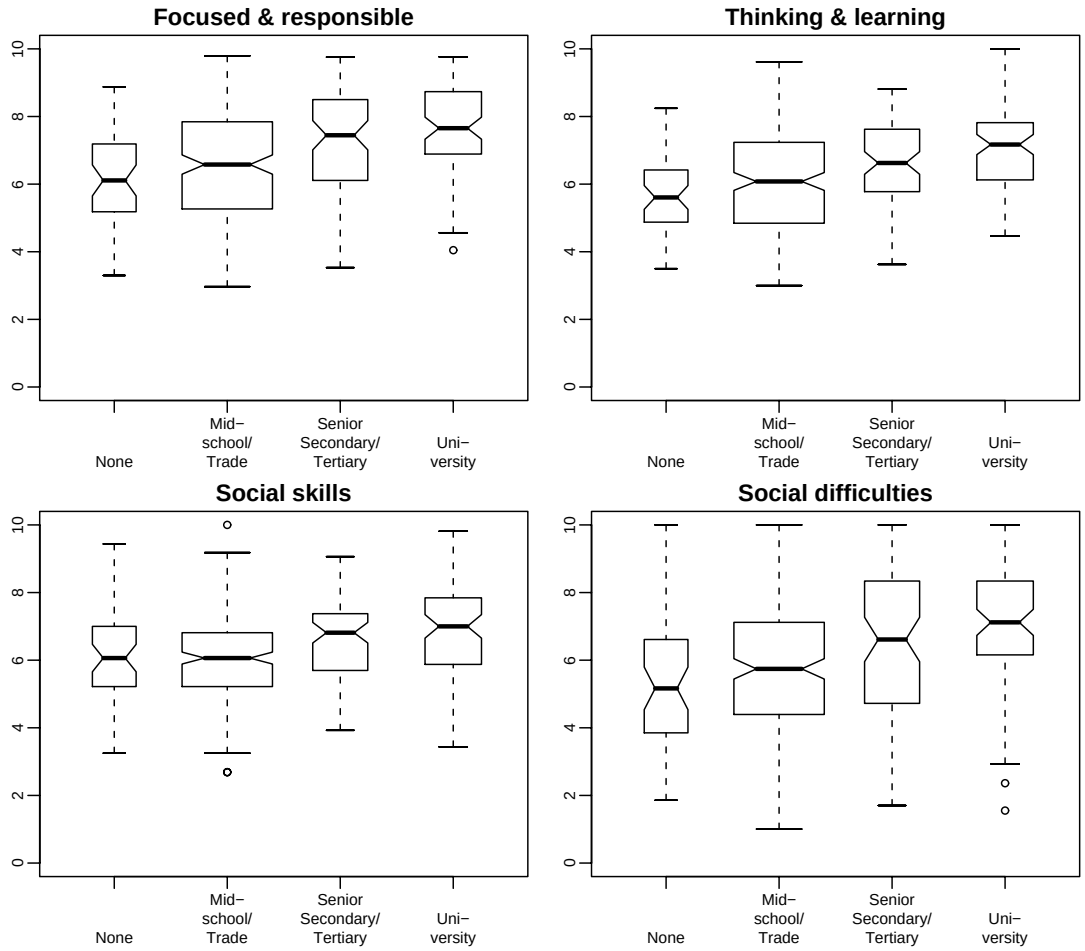


Figure 9 shows that:

- There is a gradient across the groups for all competencies, but it is weakest for social skills.
- The highest and lowest scores in each group are relatively similar
- The scores in all attitudinal competencies of those with university-qualified mothers are likely to be significantly higher than the scores of those whose mothers had no formal qualifications.

Figure 10 ***Cognitive competencies for measured 1–10 scales for family income groups***

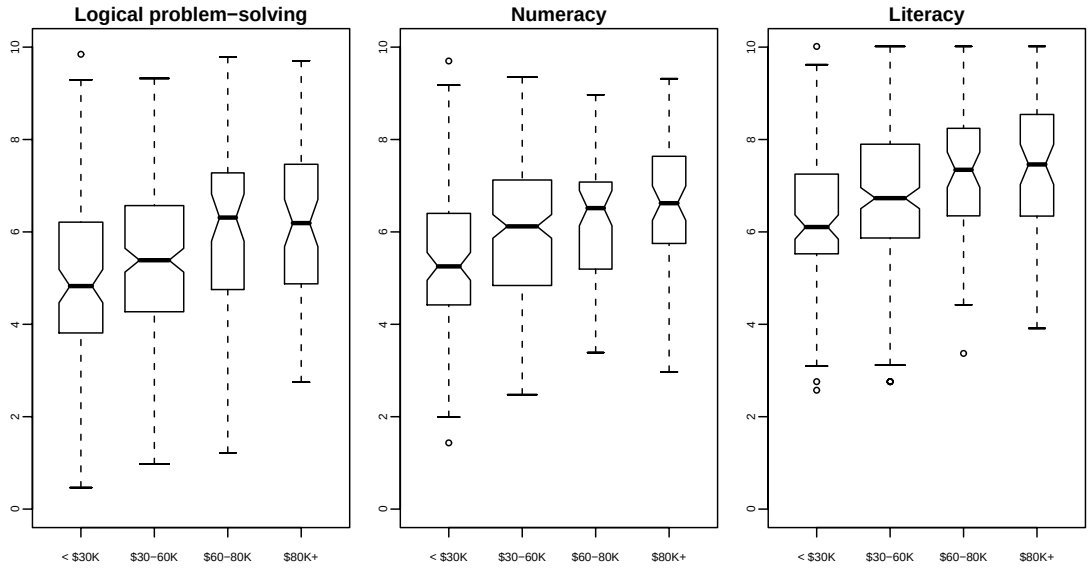


Figure 10 shows that:

- In all three competencies, there are few differences in the highest score in each of the groups, and relatively slight differences in the lowest scores, too
- The average score tended to increase with increasing family income, but the largest "jump" between groups was between the two lowest income groups. The differences between the two highest income groups were slight.

Figure 11 ***Attitudinal competencies for measured 1–10 scales for family income groups***

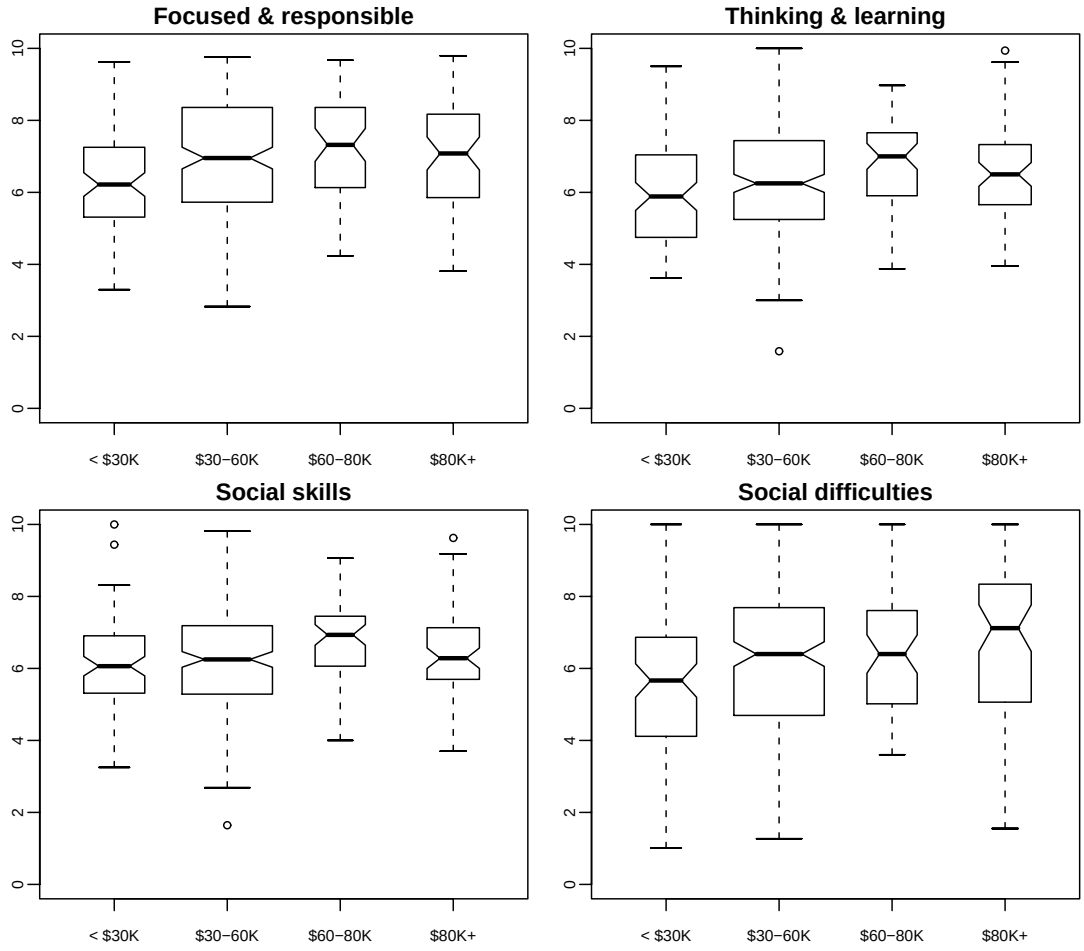


Figure 11 shows that:

- There is a slight gradient across the lowest three income groups, but, as we also found at age 14, the trend does not always continue to the highest income group
- The between-group differences are least marked for social skills.

It appears possible that there is not so much a possible advantage from living in a high-income family as there is a disadvantage from living in a low-income family. This disadvantage is most marked in the cognitive competencies.

Figure 12 ***Cognitive competencies for measured 1–10 scales for gender***

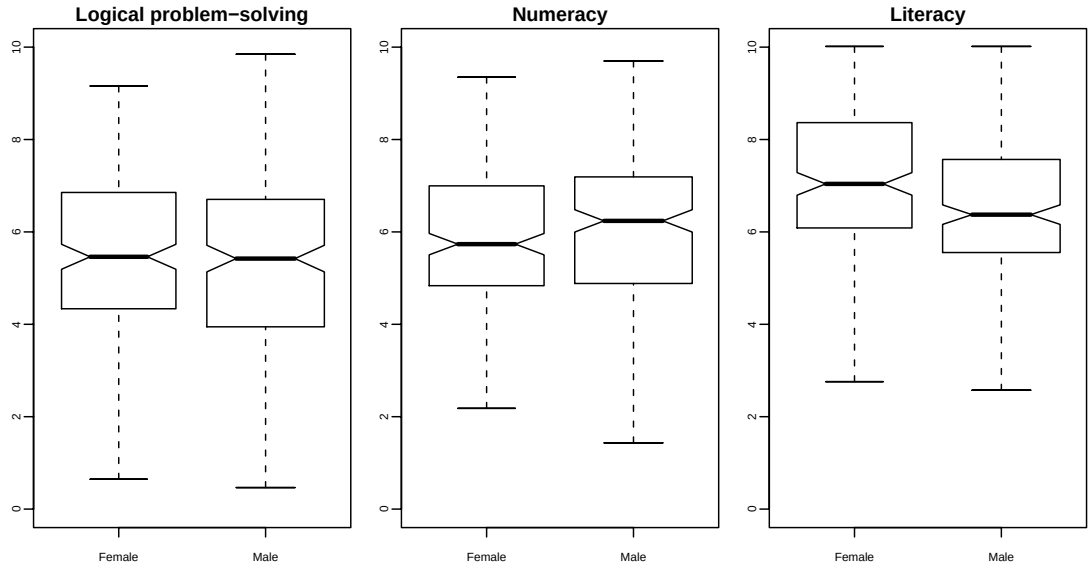
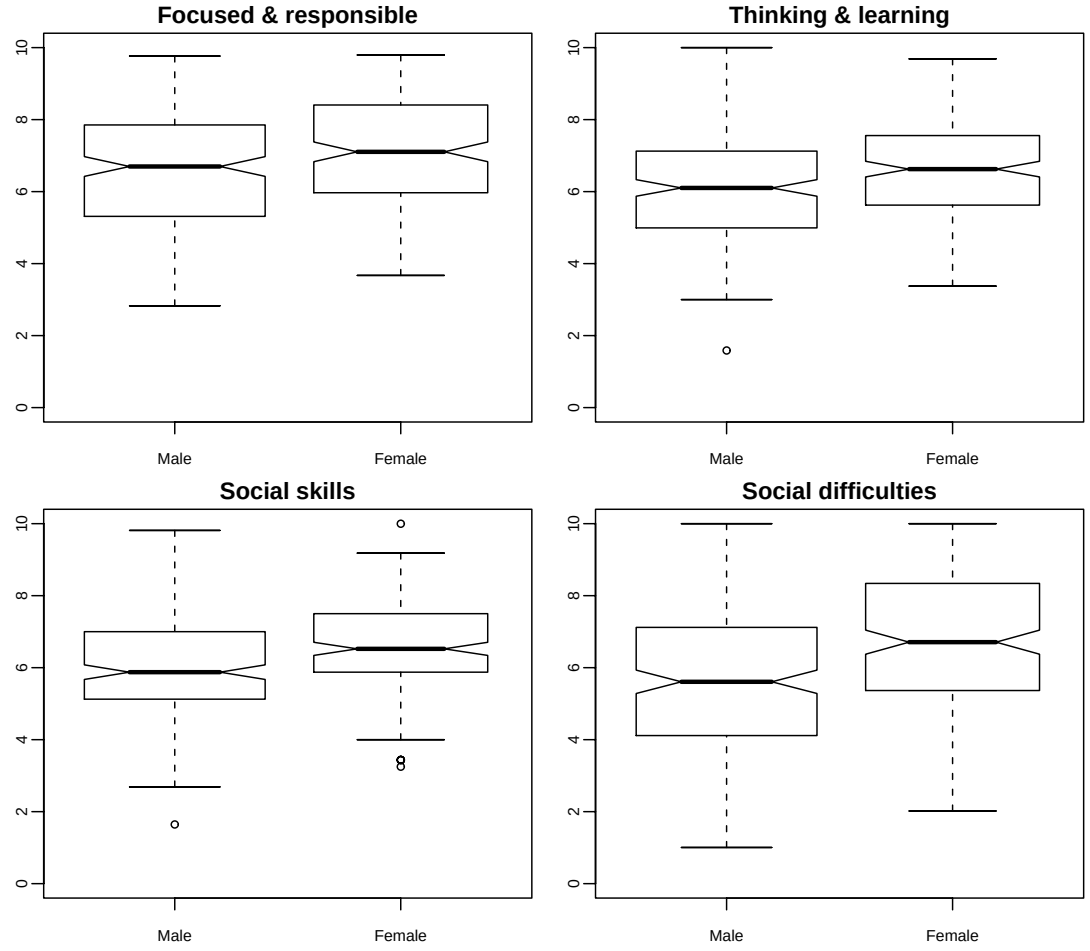


Figure 13 ***Attitudinal competencies for measured 1–10 scales for gender***



Figures 12 and 13 show that:

- There were no gender differences in the *logical problem-solving* scores, but the males achieved higher scores in *numeracy*, whereas the females achieved higher scores in *literacy* (as they had at age 14)
- The females achieved higher scores in all the attitudinal competencies which, in the case of *social difficulties*, indicated that the males were more at risk.

Figure 14 ***Cognitive competencies for measured 1–10 scales for ethnic groups***

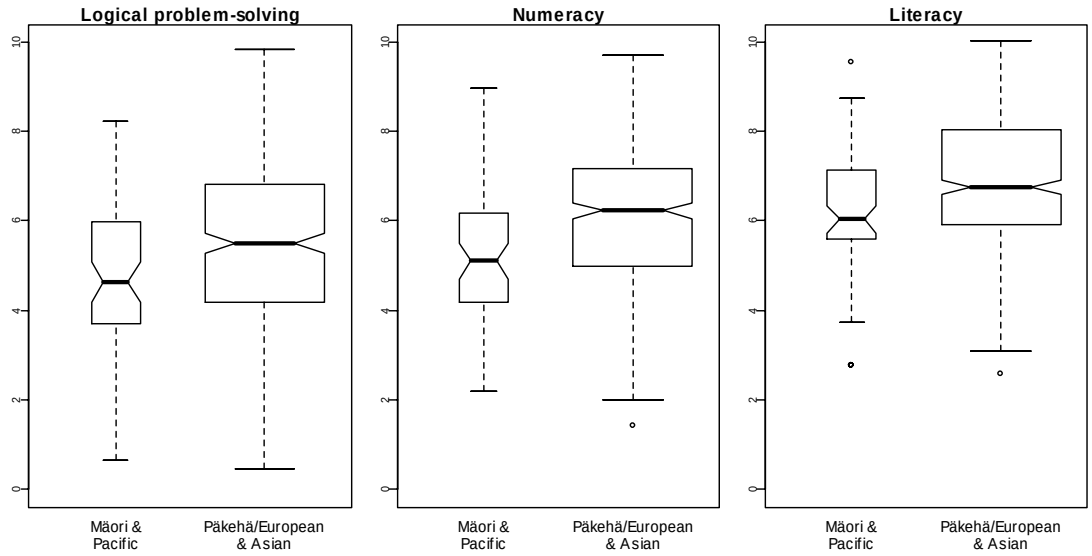
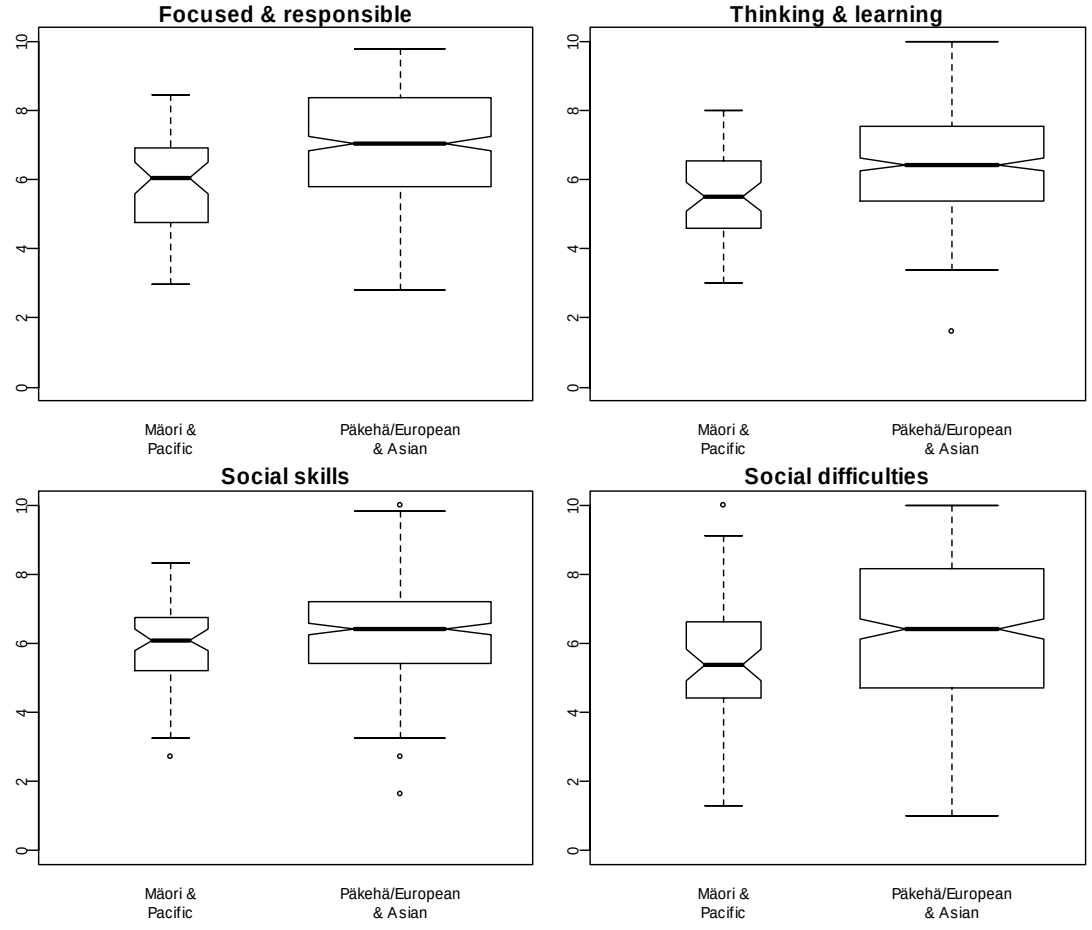


Figure 15 ***Attitudinal competencies for measured 1–10 scales for ethnic groups***



Figures 14 and 15 show that:

- Päkehä/Asian young people achieved higher scores in all the cognitive competencies than Mäori and Pacific young people; this difference was most marked for *numeracy*
- Päkehä/Asian young people also achieved higher scores in all the attitudinal competencies than Mäori and Pacific young people, with the possible exception of *social skills*.

We have now had an overview look at the differing proportions in each group that achieved above-average scores where the groups were defined by social characteristics, and have seen that there were marked differences across some social groups for some scores. We've looked also at the differences in actual score. We next do more formal tests for differences in score across the different groups.

Models of the contributions of social characteristics to each competency

We used analysis of variance[12] (ANOVA) to see how well the social characteristics explain the variability in competency score, and to see which levels of these variables are associated with higher (or lower) scores.

The models we fitted included all four social characteristics (maternal qualifications, family income at age 5, ethnic group, and gender). This is because there are associations between the variables: lower levels of maternal qualification are often (but not always) associated with lower levels of income, for instance. This allows us to see the effect of each variable after accounting for the other variables in the model. There are many other variables that would also affect the competency scores, but for the moment these are ignored. The effects of some of them are investigated in later reports in the age-16 *Competent Children, Competent Learners* series.

The models were fitted using a corner-point parameterisation, which means that all the estimates produced in the model-fitting process give the difference between the reference group for that characteristic and each of the other groups. The reference group was the male Mäori or Pacific students with mothers with no education, and an age-5 family income of under $30,000. The differences given in the tables for each variable are the estimated value while holding all other variables constant (or controlling for the effect of the other variables).

We have not included interactions in the models, which is consistent with previous reports, nor have we included a corresponding age-5 competency, which would have removed some differences in individual ability.

Numeracy

Maternal qualification was the most important predictor of the *numeracy* score; there were indicative differences for ethnic group and family income as well, where the only significant differences were between the lowest and highest income groups. A model including all four social variables accounted for 15 percent of the variability (adjusted[13] R^2) in the *numeracy* scores.

---

[12] We used the `lm` function in the statistical package R (R Development Core Team, 2006) to conduct the tests. Note that ANOVA is the same as linear regression using dummy variables.

[13] The adjusted R^2 takes into account the number of parameters in the model, and is typically lower than the unadjusted R^2, which in this case is 17 percent.

In Table 5, the estimate of 0.53 in the fourth row of the body of the table shows that the average *numeracy* score of a student whose mother had mid-secondary school/tertiary school qualifications was 0.53 on the 1–10 scale higher than the average score of a student whose mother had no qualifications (with the effect of the other variables held constant). The estimate of 0.85 in the second to last row of the table shows that the average *numeracy* score of a student whose mother held a university qualification was 0.85 higher than that of a student whose mother had mid-secondary school/trade qualifications (with the effect of the other variables held constant). Average scores of students whose mothers had university qualifications and who had a family income of over $80,000 were 1.38 + 0.67 = 2.05 higher than those of students whose mothers had no qualifications and who had a family income of under $30,000 (so long as there is no interaction between maternal qualifications and family income).

Table 5 ***Parameter estimates for age-16 numeracy score***

Parameter	***Estimate***	***Standard error***	***p-value[a]***	***Effect size[b] (%)***
Intercept (mean for reference group)	4.72	0.24		
Maternal qualifications			< 0.0001	8.3
None (reference)	0.00			
Mid-secondary school/Trade	0.53	0.21	0.0116	
Senior secondary school/Tertiary	0.82	0.25	0.0011	
University	1.38	0.25	< 0.0001	
Family income at age 5[c]			0.0588	2.1
Under $30,000 (reference)	0.00			
$30,000 to $60,000	0.34	0.17	0.0448	
$60,000 to $80,000	0.43	0.23	0.0588	
Over $80,000	0.67	0.23	0.0037	
Ethnicity[c]			0.0719	1.2
Mäori/Pacific (reference)	0.00			
Päkehä/Asian	0.45	0.20	0.0220	
Gender			0.1149	0.6
Male (reference)	0.00			
Female	-0.21	0.13	0.1149	
Other significant differences were between:				
Mid-secondary school/Trade & University	0.85	0.18	< 0.0001	
Senior secondary/Tertiary & University	0.56	0.22	0.0106	

[a] The *p*-values next to the names of the social variables are those for the *F*-test as to whether the variable adds significantly to the model when fitted last. The other *p*-values are those for a *t*-test of whether the parameter estimated is zero. If the effect estimate is significant, this means that the mean for that (marginal) group is significantly higher than the mean for the reference group. Between-group *t*-tests for other (non-reference) groups are given at the bottom of the table.

[b] Partial η^2, giving the percentage of variability accounted for by the social characteristic.

[c] For these variables, and maternal qualifications (but the effect is less obvious in that case), these data were not available for some students. One reason why there are individually significant differences between groups, but the overall effect is not significant, is that there were no differences between the small "unknown" group (not shown in the table above) and the reference group.

The pattern we see in Table 5 is consistent with the patterns seen in the preceding sections, when we had a more empirical look at the *numeracy* score. Maternal qualifications explained the most (about 8 percent) of the variability in *numeracy* score. There were statistically significant differences between:

- each of the groups defined by the level of maternal qualifications other than between the two groups with mid-secondary school/trade or senior secondary/tertiary qualifications
- those whose family income was under $30,000 and those whose income was over $80,000 (and there were indicative differences between the lowest group and the two middle groups) and
- the two ethnic groups.

Gender differences are not statistically significant in the context of these other social variables.

Literacy

Maternal qualifications and gender were the most important predictors of *literacy* score accounting for about 6 and 7 percent of the variability in the score, respectively. The model containing all four social variables accounted for 16 percent of the variability in *literacy* score (Table 6).

Table 6 ***Parameter estimates for age-16 literacy score***

Parameter	***Estimate***	***Standard error***	***p-value[a]***	***Effect size[b] (%)***
Intercept (mean for reference group)	5.32	0.25		
Maternal qualifications			< 0.0001	7.1
None (reference)	0.00			
Mid-secondary school/Trade	0.34	0.21	0.1073	
Senior secondary school/Tertiary	0.76	0.25	0.0027	
University	1.20	0.26	< 0.0001	
Family income at age 5[c]			0.1112	1.8
Under $30,000 (reference)	0.00			
$30,000 to $60,000	0.27	0.17	0.1210	
$60,000 to $80,000	0.48	0.23	0.0411	
Over $80,000	0.57	0.23	0.0146	
Ethnicity[c]			0.0609	1.3
Mäori/Pacific (reference)	0.00			
Päkehä/Asian	0.47	0.20	0.0190	
Gender			< 0.0001	5.9
Male (reference)	0.00			
Female	0.67	0.13	< 0.0001	
Other significant differences were between:				
Mid-secondary school/Trade & Senior secondary/Tertiary	0.42	0.18	0.0236	
Mid-secondary school/Trade & University	0.85	0.19	< 0.0001	

[a] The *p*-values next to the names of the social variables are those for the *F*-test as to whether the variable adds significantly to the model when fitted last. The other *p*-values are those for a *t*-test of whether the parameter estimated is zero. If the effect estimate is significant, this means that the mean for that (marginal) group is significantly higher than the mean for the reference group. Between-group *t*-tests for other (non-reference) groups are given at the bottom of the table.

[b] Partial η^2, giving the percentage of variability accounted for by the social characteristic.

[c] For these variables, and maternal qualifications (but the effect is less obvious in that case), these data were not available for some students. One reason why there are individually significant differences between groups, but the overall effect is not significant, is that there were no differences between the small "unknown" group (not shown in the table above) and the reference group.

There were statistically significant differences between:

- those whose mothers had no formal qualifications and those whose mothers had a senior secondary school/tertiary or university qualification
- those whose mothers had mid-secondary school/trade qualifications and those whose mothers had a senior secondary school/tertiary or university qualification
- males and females: on average females scored 0.67 (on the 1–10 scale) more than males (holding other variables constant) and
- Päkehä/European and Asian young people and Mäori and Pacific young people.

Logical problem-solving

Maternal qualification was the only real predictor of *logical problem-solving* score, accounting for about 11 percent of the variability in the score. The model explained 14 percent of the variability in *logical problem-solving* score (Table 7).

Table 7 ***Parameter estimates for age-16 logical problem-solving score***

Parameter	***Estimate***	***Standard error***	***p-value[a]***	***Effect size[b] (%)***
Intercept (mean for reference group)	4.10	0.29		
Maternal qualifications			< 0.0001	11.3
None (reference)	0.00			
Mid-secondary school/Trade	0.73	0.25	0.0038	
Senior secondary school/Tertiary	1.37	0.30	< 0.0001	
University	1.86	0.30	< 0.0001	
Family income at age 5[c]			0.2295	1.3
Under $30,000 (reference)	0.00			
$30,000 to $60,000	0.14	0.20	0.4966	
$60,000 to $80,000	0.44	0.27	0.1048	
Over $80,000	0.57	0.27	0.0393	
Ethnicity			0.5248	0.3
Mäori/Pacific (reference)	0.00			
Päkehä/Asian	0.19	0.23	0.4257	
Gender			0.7231	< 0.01
Male (reference)	0.00			
Female	0.06	0.16	0.7231	
Other significant differences were between:				
Mid-secondary school/Trade & Senior secondary/Tertiary	0.64	0.21	0.0030	
Mid-secondary school/Trade & University	1.13	0.22	< 0.0001	

[a] The *p*-values next to the names of the social variables are those for the *F*-test as to whether the variable adds significantly to the model when fitted last. The other *p*-values are those for a *t*-test of whether the parameter estimated is zero. If the effect estimate is significant, this means that the mean for that (marginal) group is significantly higher than the mean for the reference group. Between-group *t*-tests for other (non-reference) groups are given at the bottom of the table.

[b] Partial η^2, giving the percentage of variability accounted for by the social characteristic.

[c] For these variables, and maternal qualifications (but the effect is less obvious in that case), these data were not available for some students.

There were statistically significant differences between:

- those whose mothers had no formal qualifications and those whose mothers had any level of qualification
- those whose mothers had mid-secondary school/trade qualifications and those whose mothers had a senior secondary school/tertiary or university qualification.

Thinking & learning

Maternal qualifications, ethnicity, and gender made significant contributions to the model, accounting for 6, 3, and 3 percent of the variability, respectively. The model as a whole accounted for 13 percent of the variability in the *thinking & learning* score (Table 8).

Table 8 ***Parameter estimates for age-16 thinking & learning score***

Parameter	***Estimate***	***Standard error***	***p-value[a]***	***Effect size[b] (%)***
Intercept (mean for reference group)	4.98	0.27		
Maternal qualifications			< 0.0001	6.1
None (reference)	0.00			
Mid-secondary school/Trade	0.15	0.22	0.5139	
Senior secondary school/Tertiary	0.62	0.26	0.0175	
University	0.98	0.27	0.0003	
Family income at age 5[c]			0.1917	1.5
Under $30,000 (reference)	0.00			
$30,000 to $60,000	0.16	0.18	0.3842	
$60,000 to $80,000	0.47	0.24	0.0509	
Over $80,000	0.12	0.24	0.6117	
Ethnicity[c]			0.0029	3.0
Mäori/Pacific (reference)	0.00			
Päkehä/Asian	0.70	0.20	0.0007	
Gender			0.0008	2.8
Male (reference)	0.00			
Female	0.46	0.14	0.0008	
Other significant differences were between:				
Mid-secondary school/Trade & Senior secondary/Tertiary	0.47	0.18	0.0104	
Mid-secondary school/Trade & University	0.83	0.19	< 0.0001	

[a] The *p*-values next to the names of the social variables are those for the *F*-test as to whether the variable adds significantly to the model when fitted last. The other *p*-values are those for a *t*-test of whether the parameter estimated is zero. If the effect estimate is significant, this means that the mean for that (marginal) group is significantly higher than the mean for the reference group. Between-group *t*-tests for other (non-reference) groups are given at the bottom of the table.

[b] Partial η^2, giving the percentage of variability accounted for by the social characteristic.

[c] For these variables, and maternal qualifications (but the effect is less obvious in that case), these data were not available for some students.

The social characteristics that accounted for the most variability in *thinking & learning* score were maternal qualifications (effect size of 6.1), ethnicity (effect size of 3.0), and gender (effect size of 2.8). There were statistically significant differences between:

- those whose mothers had no formal qualifications and those whose mothers had a senior secondary school/tertiary or university qualification
- those whose mothers had mid-secondary school/trade qualifications and those whose mothers had a senior secondary school/tertiary or university qualification
- males and females and
- Päkehä/European and Asian young people and Mäori and Pacific young people.

Focused & responsible

Maternal qualifications, ethnicity, and gender all made significant contributions to the model, accounting for 9, 4, and 4 percent of the variability, respectively. The model as a whole accounted for 16 percent of the variability in the *focused & responsible* score (Table 9).

Table 9 ***Parameter estimates for age-16 focused & responsible score***

Parameter	***Estimate***	***Standard error***	***p-value[a]***	***Effect size[b] (%)***
Intercept (mean for reference group)	5.14	0.28		
Maternal qualifications			< 0.0001	8.9
None (reference)	0.00			
Mid-secondary school/Trade	0.21	0.24	0.3903	
Senior secondary school/Tertiary	0.89	0.28	0.0016	
University	1.25	0.29	< 0.0001	
Family income at age 5[c]			0.3997	1.0
Under $30,000 (reference)	0.00			
$30,000 to $60,000	0.33	0.19	0.0878	
$60,000 to $80,000	0.39	0.26	0.1248	
Over $80,000	0.16	0.25	0.5284	
Ethnicity[c]			0.0005	3.9
Mäori/Pacific (reference)	0.00			
Päkehä/Asian	0.80	0.22	0.0003	
Gender			< 0.0001	4.4
Male (reference)	0.00			
Female	0.60	0.15	< 0.0001	
Other significant differences were between:				
Mid-secondary school/Trade & Senior secondary/Tertiary	0.68	0.20	0.0006	
Mid-secondary school/Trade & University	1.04	0.20	< 0.0001	

[a] The *p*-values next to the names of the social variables are those for the *F*-test as to whether the variable adds significantly to the model when fitted last. The other *p*-values are those for a *t*-test of whether the parameter estimated is zero. If the effect estimate is significant, this means that the mean for that (marginal) group is significantly higher than the mean for the reference group. Between-group *t*-tests for other (non-reference) groups are given at the bottom of the table.

[b] Partial η^2, giving the percentage of variability accounted for by the social characteristic.

[c] For these variables, and maternal qualifications (but the effect is less obvious in that case), these data were not available for some students.

There were statistically significant differences between:

- those whose mothers had no formal qualifications and those whose mothers had a senior secondary school/tertiary or university qualification
- those whose mothers had mid-secondary school/trade qualifications and those whose mothers had a senior secondary school/tertiary or university qualification
- males and females and
- Päkehä/European and Asian young people and Mäori and Pacific young people.

Social skills

Maternal qualifications and gender made significant contributions to the model, each accounting for 5 percent of the variability in score. The whole model accounted for 10 percent of the variability in the *social skills* score (Table 10).

Table 10 ***Parameter estimates for age-16 social skills score***

Parameter	***Estimate***	***Standard error***	***p-value***[a]	***Effect size***[b] ***(%)***
Intercept (mean for reference group)	5.55	0.25		
Maternal qualifications			0.0002	5.2
None (reference)	0.00			
Mid-secondary school/Trade	-0.10	0.21	0.6388	
Senior secondary school/Tertiary	0.47	0.24	0.0542	
University	0.59	0.25	0.0192	
Family income at age 5[c]			0.3929	1.0
Under $30,000 (reference)	0.00			
$30,000 to $60,000	0.07	0.17	0.6692	
$60,000 to $80,000	0.40	0.22	0.0757	
Over $80,000	0.12	0.22	0.5937	
Ethnicity[c]			0.4118	0.4
Mäori/Pacific (reference)	0.00			
Päkehä/Asian	0.23	0.19	0.2346	
Gender			< 0.0001	5.3
Male (reference)	0.00			
Female	0.59	0.13	< 0.0001	
Other significant differences were between:				
Mid-secondary school/Trade & Senior secondary/Tertiary	0.57	0.17	0.0011	
Mid-secondary school/Trade & University	0.69	0.18	0.0001	

[a] The *p*-values next to the names of the social variables are those for the *F*-test as to whether the variable adds significantly to the model when fitted last. The other *p*-values are those for a *t*-test of whether the parameter estimated is zero. If the effect estimate is significant, this means that the mean for that (marginal) group is significantly higher than the mean for the reference group. Between-group *t*-tests for other (non-reference) groups are given at the bottom of the table.

[b] Partial η^2, giving the percentage of variability accounted for by the social characteristic.

[c] For these variables, and maternal qualifications (but the effect is less obvious in that case), these data were not available for some students.

There were statistically significant differences between:

- those whose mothers had no formal qualifications and those whose mothers had a senior secondary school/tertiary or university qualification
- those whose mothers had mid-secondary school/trade qualifications and those whose mothers had a senior secondary school/tertiary or university qualification and
- males and females.

Social difficulties

Maternal qualifications and gender made a significant contribution to the model, each accounting for about 6 percent of the variability in score. The whole model accounted for 13 percent of the variability in the *social difficulties* score (Table 11). Note that a high score indicates an *absence* or low risk of difficulty.

Table 11 ***Parameter estimates for age-16 social difficulties score***

Parameter	***Estimate***	***Standard error***	***p-value[a]***	***Effect size[b] (%)***
Intercept (mean for reference group)	4.36	0.39		
Maternal qualifications			< 0.0001	6.0
None (reference)	0.00			
Mid-secondary school/Trade	0.16	0.33	0.6246	
Senior secondary school/Tertiary	1.10	0.38	0.0040	
University	1.24	0.39	0.0017	
Family income at age 5[c]			0.3514	1.1
Under $30,000 (reference)	0.00			
$30,000 to $60,000	0.44	0.26	0.0934	
$60,000 to $80,000	0.28	0.35	0.4218	
Over $80,000	0.67	0.35	0.0543	
Ethnicity[c]			0.1901	0.8
Mäori/Pacific (reference)	0.00			
Päkehä/Asian	0.54	0.30	0.0746	
Gender			< 0.0001	6.9
Male (reference)	0.00			
Female	1.04	0.20	< 0.0001	
Other significant differences were between:				
Mid-secondary school/Trade & Senior secondary/Tertiary	0.94	0.27	0.0005	
Mid-secondary school/Trade & University	1.07	0.28	0.0001	

[a] The *p*-values next to the names of the social variables are those for the *F*-test as to whether the variable adds significantly to the model when fitted last. The other *p*-values are those for a *t*-test of whether the parameter estimated is zero. If the effect estimate is significant, this means that the mean for that (marginal) group is significantly higher than the mean for the reference group. Between-group *t*-tests for other (non-reference) groups are given at the bottom of the table.

[b] Partial η^2, giving the percentage of variability accounted for by the social characteristic.

[c] For these variables, and maternal qualifications (but the effect is less obvious in that case), these data were not available for some students.

There were statistically significant differences between:

- those whose mothers had no formal qualifications and those whose mothers had a senior secondary school/tertiary or university qualification
- those whose mothers had mid-secondary school/trade qualifications and those whose mothers had a senior secondary school/tertiary or university qualification and
- males and females.

Which social characteristics contribute most to differences in age-16 scores?

Table 12 summarises which social variables were significant predictors of which competency scores, and the proportion of variance accounted for by each of them (the effect sizes) and the model as a whole.

Table 12 ***Social variables' contribution to the variance in age-16 scores***

Competency					***Proportion of variance accounted for by social characteristics %***
Numeracy	Maternal qualification (8.3)	Family income (2.1)		Ethnicity (1.2)	15
Logical problem-solving	Maternal qualification (11.3)				14
Literacy	Maternal qualification (7.1)	Family income (1.8)	Gender (5.9)	Ethnicity (1.3)	16
Thinking & learning	Maternal qualification (6.1)		Gender (2.8)	Ethnicity (3.0)	13
Focused & responsible	Maternal qualification (8.9)		Gender (4.4)	Ethnicity (3.9)	16
Social skills	Maternal qualification (5.2)		Gender (5.3)		10
Social difficulties	Maternal qualification (6.0)		Gender (6.9)		13

What the tests above have shown is that maternal qualification levels, which may be associated with whether a young person is more likely to be intellectually advantaged or not, and to have had a language-rich home environment or not, are strongly associated with all the competencies. The effects of maternal qualification were least strong in social skills.

Age-5 family income, which may be associated with resources available in the home environment, the (non)deprived home environment, early childhood education, and school socioeconomic mix, is relatively strongly associated with *literacy* and *numeracy*, with increasing advantage being associated with increasing competency scores up to, but not beyond, a "comfortable middle-income" level. The advantages conferred by a higher age-5 family income are not marked in the attitudinal competencies.

Ethnicity is most strongly associated with the attitudinal competencies of *thinking & learning* and *focused & responsible*. It is not associated with *logical problem-solving*, nor with *social skills*, or *social difficulties*. In the models fitted, where ethnicity was significant, it indicated a significant effect over and above effects due to maternal qualification and age-5 income.[14]

[14] In the next report in this series (Hodgen, forthcoming), a corresponding age-5 competency was added to the model, as well as one of the ECE experience variables. In each case, once the age-5 competency was included, ethnicity no longer added significantly to the model. This could be interpreted to mean that any differences in competency between ethnic groups

Gender differences were greatest in *literacy* and the attitudinal competencies, particularly *focused & responsible* and the social competencies. Gender differences were not statistically significant in *logical problem-solving* and *numeracy*.

Comparison with the situation at age 14

In the corresponding report at age 14 we found that family income made a significant contribution to more of the models than we have found in our analysis of age-16 competency scores. Chapter 8 reports that the group of participants who withdrew from the study at age 16 were predominantly from low-income homes, and that Mäori and Pacific young people were over-represented in the group. To test whether the difference in findings between ages 14 and 16 were due to the reduced number of students from low-income families, we repeated the age-14 analysis, using age-14 competencies, on the same individuals who were included in the age-16 analysis. The results are given in Table 13, where the numbers in brackets after the social characteristics are the relevant effect sizes.

Table 13 ***Social variables' contribution to the variance in age-14 scores for age-16 participants***

Competency					***Proportion of variance accounted for by social characteristics %***
Mathematics	Maternal qualification (11.8)	Family income (3.4)		Ethnicity (2.1)	21
Logical problem-solving	Maternal qualification (12.3)				13
Reading comprehension	Maternal qualification (10.5)	Family income (4.5)	Gender (1.5)	Ethnicity (1.5)	21
Perseverance	Maternal qualification (7.6)		Gender (6.6)	Ethnicity (3.2)	16
Self-management	Maternal qualification (6.4)		Gender (7.0)	Ethnicity (2.0)	13
Social skills	Maternal qualification (6.1)		Gender (6.7)		12
Self-efficacy	Maternal qualification (6.3)		Gender (4.7)		12
Communication	Maternal qualification (9.5)		Gender (5.5)	Ethnicity (2.0)	16
Curiosity	Maternal qualification (5.3)				8

had developed, or existed, before age 5, and that any such differences had not increased between the ages of 5 and 16, once the other social characteristics were accounted for. Age-5 family income was also no longer statistically significant in these models, with similar implications. However, maternal qualifications and gender remained significant in the models.

The pattern we see in this table is the same as that in Table 12 (age 16). For the participants still in the study at age 16, family income at age 5 adds significantly to *literacy* and *numeracy* scores, but no others. It also accounts for slightly less variation in these scores at age 16 than at age 14, as does maternal qualification. This suggests that the greater role played by family income found at age 14 reflected a higher proportion of low-income participants at that age in the sample. That then suggests that if our sample was truly representative of the New Zealand student population, family income might play more of a role in differences in student performance.

In general, at both age 14 and age 16, maternal qualification accounts for more of the variation in scores than any of the other characteristics; sometimes with the attitudinal competencies the amounts of variation accounted for by maternal qualifications and gender are approximately equal; ethnicity accounts for less variation than the other social characteristics; and family income accounts for statistically significant amounts of variation in *literacy* and *numeracy* only.

Summary

We found that the four social characteristics (maternal qualifications, age-5 family income, gender, and ethnicity) together explained 10–15 percent of the variability in the three cognitive and four attitudinal competency scores.

As we had found before, maternal qualifications was the most important (representing a number of aspects, such as ability, a language- and stimulation-rich environment, value being placed on a good education, and schools being chosen with this in mind), accounting for 5–11 percent of the variability in the competency scores.

Where there was a gender effect (there was none in *numeracy* and *logical problem-solving*), it was next most important, accounting for 3–7 percent of the variability in score. Gender accounted for more of the variation in *literacy* and *social difficulty* scores than in any of the other scores. Neither ethnicity nor age-5 family income were particularly important, accounting for only 1–3 percent of the variability (and this contribution was no longer significant when an equivalent age-5 score was added to the model (Hodgen, forthcoming)).