NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH

Multilevel Analysis of PIRLS Data for England

Ian Schagen
i.schagen@nfer.ac.uk

Paper presented at the International Research Conference (IRC-2004)
University of Cyprus, Lefkosia
11-13 May 2004

Multilevel Analysis of PIRLS Data for England

	Ian Schagen
	National Foundation for Educational Research
	The Mere, Upton Park, Slough SL1 2DQ, UK

Abstract

The data collected as part of the PIRLS study in England yielded a number of measures which were used in this analysis: internationally-derived scales for pupils' cognitive outcomes (3 scales), nationally-derived factor scores for pupils' attitudes to reading, pupil background information and school background and attitude information. The aim of this analysis was to investigate factors at the school and pupil levels which might be associated with the various international scales, and to see which were apparently statistically significant. In addition, some analysis was carried out on factors related to the pupil attitude scales.

Keywords

International, reading, multilevel.

Introduction

The aim of this study was to explore the relationships between pupils' attitudes to reading, as well as school and pupil background characteristics, to those pupils' attainment in the PIRLS reading tests in England (see Twist et al, 2003, for a fuller description of PIRLS results for England). To do this, the following stages of analysis were carried out:

1. Derive pupil-level attitude measures from the pupil questionnaire data.
2. Derive school-level aggregated measures from the school questionnaire data.
3. Combine the above with pupil and school background data into a series of multilevel models to explore the relationships between these measures and pupil attainment as measured by the PIRLS reading tests.
4. Additionally, run multilevel models to explore the relationships between pupil attitudes to reading and other background factors.

Detailed results of this analysis are presented in this paper, both graphically and numerically, and some conclusions are presented, together with suggestions for further work in this area.

Development of Student Attitude Scales

The student questionnaire data was analysed to develop a set of student attitude scales., using the following methodology:

1. Identify groups of attitude items which seem to relate together, using exploratory factory analysis;
2. Rescale items to a new metric ('Never' = 0, 'Every day' = 30; or 'Disagree a lot' = -20, 'Agree a lot' = 20);
3. Calculate attitude scale score as mean of item responses.

Although simple, this approach has the advantage that it is possible to compare each scale's mean value with the other scales, and hence evaluate the relative strength of feeling about each.

The exploratory factor analysis produced two plausible solutions, one with 3 factors and one with 7. Scales based on the former are known as 'major', while those based on the 7-factor solution have been called 'minor'.

Tables 1 and 2 shows the major and minor national scales developed, the items on which they were based, their mean values, their reliability indices, and the correlations with the overall standardised test score. The reliability value (based on Cronbach's alpha) is an indicator of the extent to which the items which make up each scale are mutually correlated, and hence measuring essentially the same construct. Values close to 1.0 are perfect, and values around 0 would imply no mutual relationship. The reliability index tends to increase with more items in the scale. Taking this into account, it seems that most of the scales have acceptable levels of reliability.

Table 1: National Attitude Scales (Major)

Scale Name	**Scale Items**	**Description**	**Mean Value**	**Number of Items**	**Reliability of Scale**	**Correlation with score**
STMAJFC1	Q3a-d,f, q4c, q4f, q5, q7b,c,e, q8a-h, q11bb, q12b	Reading activities	11.66	21	0.84	-0.15
STMAJFC2	Q3c-e, q4b, q7d, q7f, q12a-d,f, q13a-c	Reading enjoyment & confidence	11.47	14	0.79	0.38
STMAJFC3	Q11aa,ac, q11ba-bd	Use of computers	14.30	6	0.73	-0.05

Table 2: National Attitude Scales (Minor)

Scale Name	Scale Items	Description	Mean Value	Number of Items	Reliability of Scale	Correlation with score
STMINFC1	Q3d,e, q4b, q7d, q10, q12a-d,f	Reading enjoyment	11.58	10	0.78	0.31
STMINFC2	Q7b,c,e, q8a-h	Reading activities in class	11.24	11	0.77	-0.18
STMINFC3	Q3a-d,f, q4a,c,d,f,g	Reading activities at home	14.76	10	0.70	-0.11
STMINFC4	Q11aa,ac, q11ba-bd	Use of computers	14.30	6	0.73	-0.05
STMINFC5	Q13a-c	Reading confidence	7.53	3	0.53	0.40
STMINFC6	Q3g, q6	TV	24.62	2	0.42	0.09
STMINFC7	Q3c, q8d, q12b	Talking about reading	6.41	3	0.68	0.04

Development of School Questionnaire Scales

As part of the national analysis of the PIRLS data for England, the school questionnaire data was analysed to develop a set of scales.
The exploratory factor analysis was carried out on two groups of variables, labelled 'school background' and 'school policy'. The former were assumed to be those things over which the school had little or no control, while the latter comprised those elements which the school's policies could affect. The exploratory factor analysis suggested 3 'background' and 4 'policy' factors.

Tables 3 and 4 shows the scales developed, the items on which they were based, their mean values, their reliability indices, and the correlations with the average overall standardised test score for the students in the school (if significant). From these results, it seems that most of the scales have acceptable levels of reliability.

Table 3: School Background Scales

Scale Name	Scale Items	Description	Mean Value	Number of Items	Reliability of Scale	Correlation with score
SCHBGFC1	Q14.1-5	Attainment at Year 1	19.56	5	0.95	0.25
SCHBGFC2	Q4, q5, q7, q8.1-4, q9.1-3, q24.1, q26.4, q27	Disadvantaged background	-1.69	13	0.82	-0.58

SCHBGFC3	Q6.2-6	Closeness to cultural & educational facilities	18.54	5	0.64	-

Table 4: School Policy Scales

Scale Name	Scale Items	Description	Mean Value	Number of Items	Reliability of Scale	Correlation with score
SCHPLFC1	Q19.1-12	Speed of introducing reading elements	29.45	11	0.97	-
SCHPLFC2	Q15.1-3, q25.2,5,6	Emphasis on English & parental involvement	19.61	6	0.57	0.15*
SCHPLFC3	Q13.4,5	Parents & children influence curriculum	5.71	2	0.83	-
SCHPLFC4	Q19.9,11,12 versus q19.1,3,4	Higher v. lower reading skills	-4.69	6	0.27	-

(* - correlation significant at 10% level)

Multilevel Analysis of PIRLS Data for England

The following types of data were available for the analysis of the international study outcomes in England:

- Internationally-derived scales for pupils' cognitive outcomes (3 scales);
- Nationally-derived factor scores for pupils' attitudes to reading (7 scales);
- Pupil background information;
- School background and attitude information.

Information was also available about teachers and parents. At a later stage, scales derived from these instruments may also be included in the modelling. The aim of this analysis was to investigate factors at the school and pupil levels which might be associated with the various international scales, and to see which were apparently statistically significant. In addition, some analysis was carried out on factors related to the 7 pupil attitude scales.

Multilevel modelling is a development of a common statistical technique known as 'regression analysis'. This is a technique for finding a straight-line relationship which

allows us to predict the values of some measure of interest ('dependent variable') given the values of one or more related measures. For example, we may wish to predict schools' average test performance given some background factors, such as free school meals and school size (these are sometimes called 'independent variables').

Multilevel modelling takes account of data which is grouped into similar clusters at different levels. For example, individual pupils are grouped into schools, and those schools are grouped within local education authorities (LEAs). There may be more in common between pupils within the same school than with other schools, and there may be elements of similarity between different schools in the same LEA. Multilevel modelling allows us to take account of this hierarchical structure of the data and produce more accurate predictions, as well as estimates of the differences between pupils, between schools, and between LEAs. Goldstein (2003) gives a much fuller description of multilevel modelling and its application to educational data.

Ideally, it would be good to have an indicator of pupils' prior attainment, for example at the start of key stage 2. This data is not currently available, but it is hoped at a later stage to have access to pupils' national curriculum results at the end of key stage 1. The analysis reported here is not 'value-added' in any sense.

The models fitted to the data incorporated three levels:

1. LEA;
2. School;
3. Pupil.

Thus, there are assumed to be variations between LEAs in their average scores, between schools in the same LEA, and within each school there are almost bound to be variations between pupils in their attitudes and cognitive scores. The sizes of these variations at each level of the model are measured in terms of 'random variances', and the relative sizes of these will be of some interest.

For each outcome measure (international scale), the fitting process was carried out in three stages:

1. The 'base case', with no background variables;
2. Controlling for pupil-level background variables, removing those which were clearly not significant.
3. Including school-level variables in the final model, once more removing those which were clearly not significant.

Table 5 is a full listing of each variable included in the multilevel analysis, including a brief description.

Results of Multilevel Analysis – Relationships with Background Variables

In technical language, the multilevel model results comprise the random variances at each level at each stage of model fitting, plus the coefficients of the background variables in the 'full model'. From estimated standard errors we may deduce whether

or not variances or coefficients are statistically significant at the 5% level, as well as 95% confidence intervals for each parameter.

These results may not be easy to interpret for all readers. To aid in interpretation, therefore, the coefficients which express the estimated relationships between the scales and each of the background variables have been converted into 'normalised coefficients' which represent the 'strength' of each relationship as a percentage, and which allow the different variables to be compared in terms of their apparent influence on the outcome, when all other variables are simultaneously taken into account.

Normalised coefficients are plotted in Figures 1 to 3, for each of the international scales. For each variable, the estimated normalised coefficient is plotted as a diamond, with a vertical line indicating the 95% confidence interval for the estimate. Any variable whose line intersects the horizontal zero axis can be regarded as not statistically significant (at the 5% level). Positive values imply a positive relationship with the international scale outcome; negative values imply that scale values tend to decrease with higher values of the given background variable. Table 6 is a summary of the results for all the test outcomes, in terms of their normalised coefficients. Only statistically significant values are shown in this table.

Figure 1: Normalised coefficients from Multilevel Model fitted to Total Scores

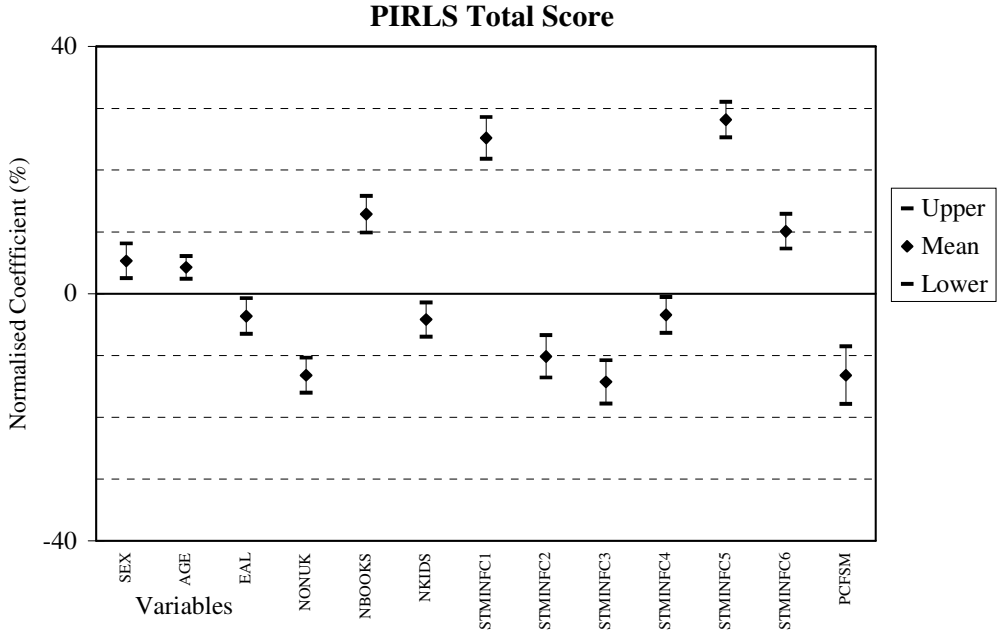


Figure 2: Normalised coefficients from Multilevel Model fitted to Information Scores

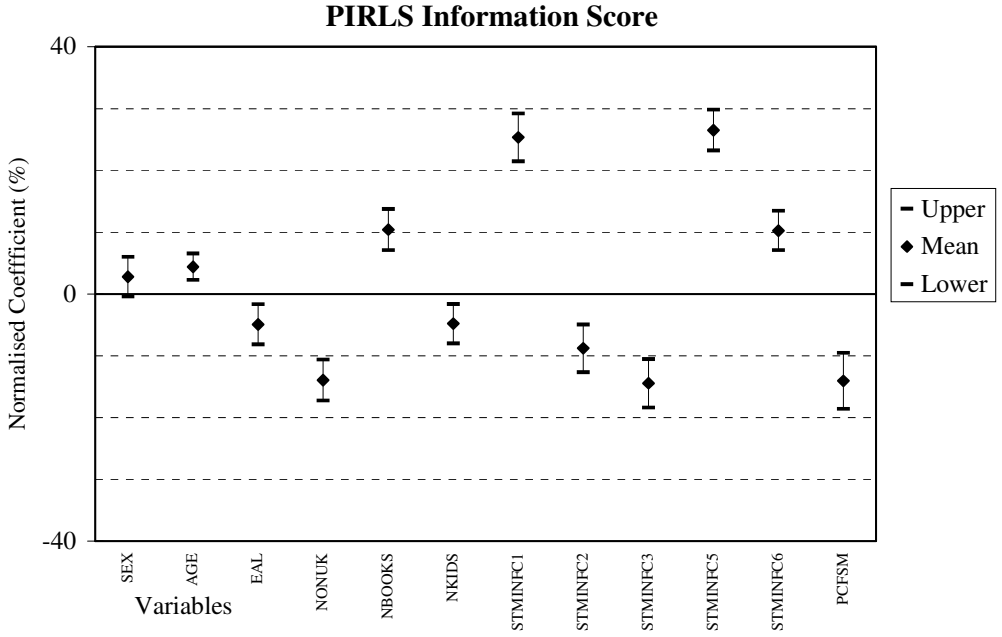
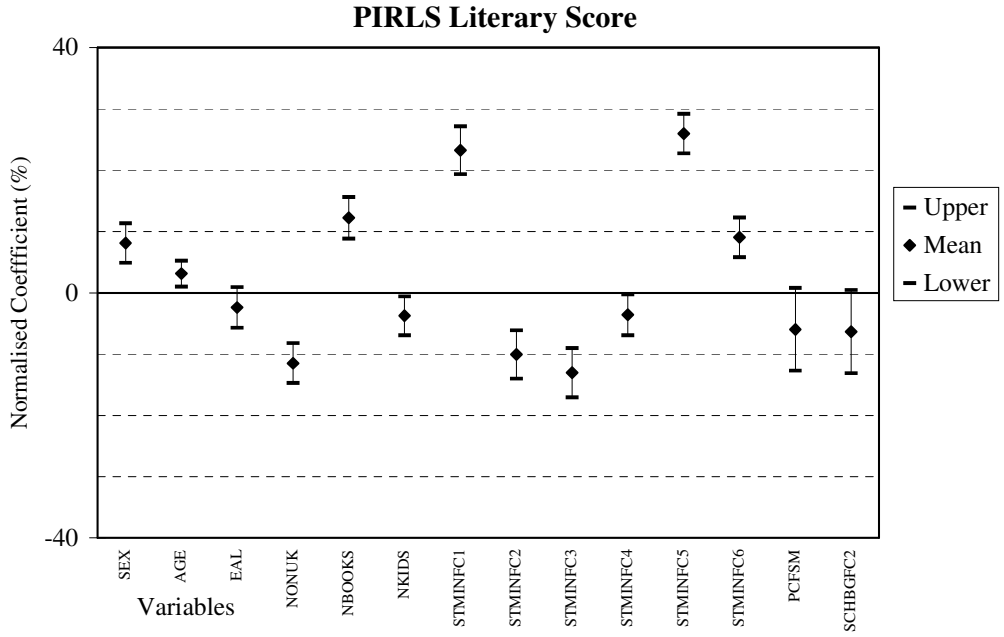


Figure 3: Normalised coefficients from Multilevel Model fitted to Literary Scores



The above analyses linked pupils' test scores to other background factors. It is clear that there are strong links between scores and attitudes, but that the school-level factors add little or nothing to the model, once pupil factors have been controlled for.

The next stage of the analysis was intended to explore the possibility of a link between school factors and pupil attitudes. Each of the 7 attitude factors was analysed separately, and the results in terms of normalised coefficients are shown in Table 7. These results are also shown graphically in Figures 4 to 10.

Figure 4: Normalised coefficients from Multilevel Model fitted to Reading Enjoyment

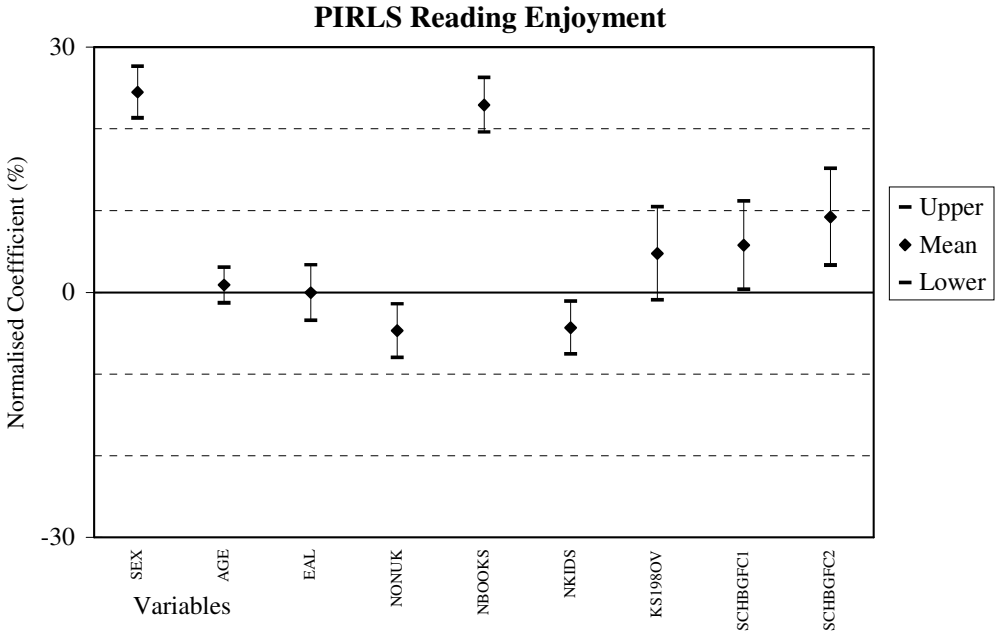


Figure 5: Normalised coefficients from Multilevel Model fitted to Reading Activities in Class

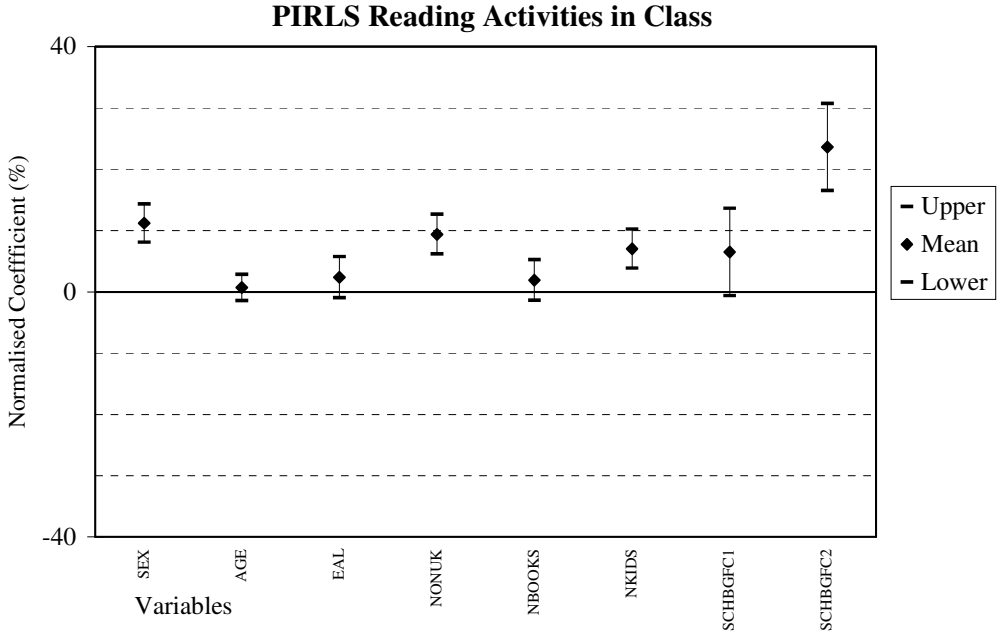


Figure 6: Normalised coefficients from Multilevel Model fitted to Reading Activities at Home

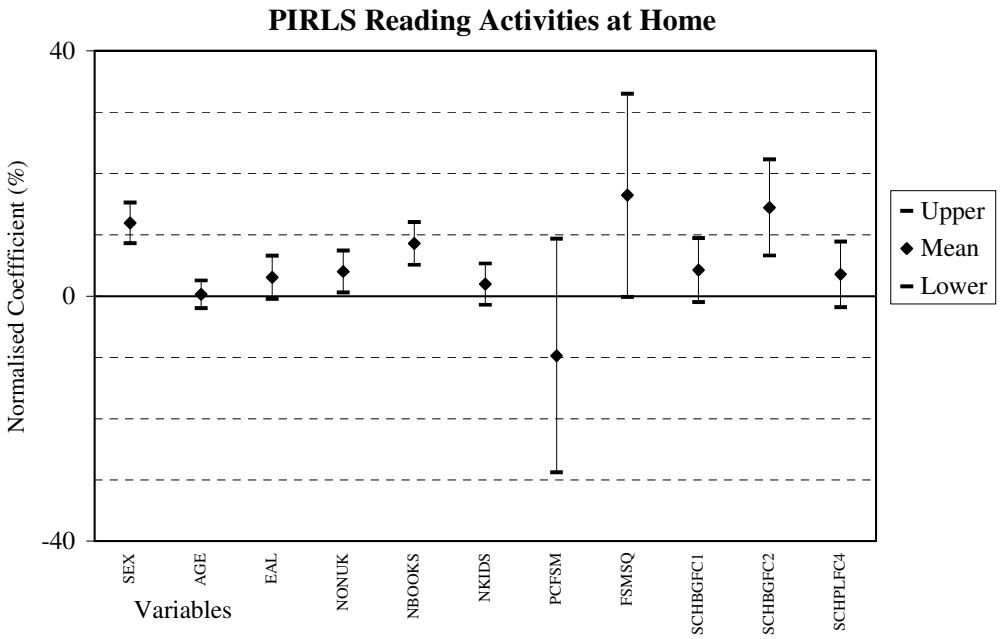


Figure 7: Normalised coefficients from Multilevel Model fitted to Use of Computers

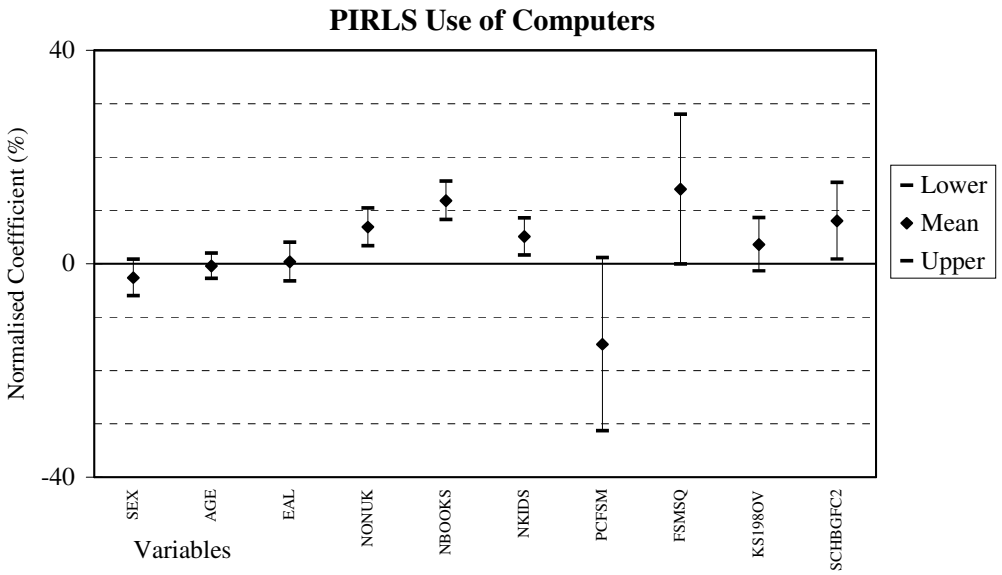

PIRLS Use of Computers
40
0
-40
Normalised Coeffficient (%)
SEX
AGE
EAL
NONUK
NBOOKS
NKIDS
PCFSM
FSMSQ
KS198OV
SCHBGFC2
Variables
Lower
Mean
Upper

Figure 8: Normalised coefficients from Multilevel Model fitted to Reading Confidence

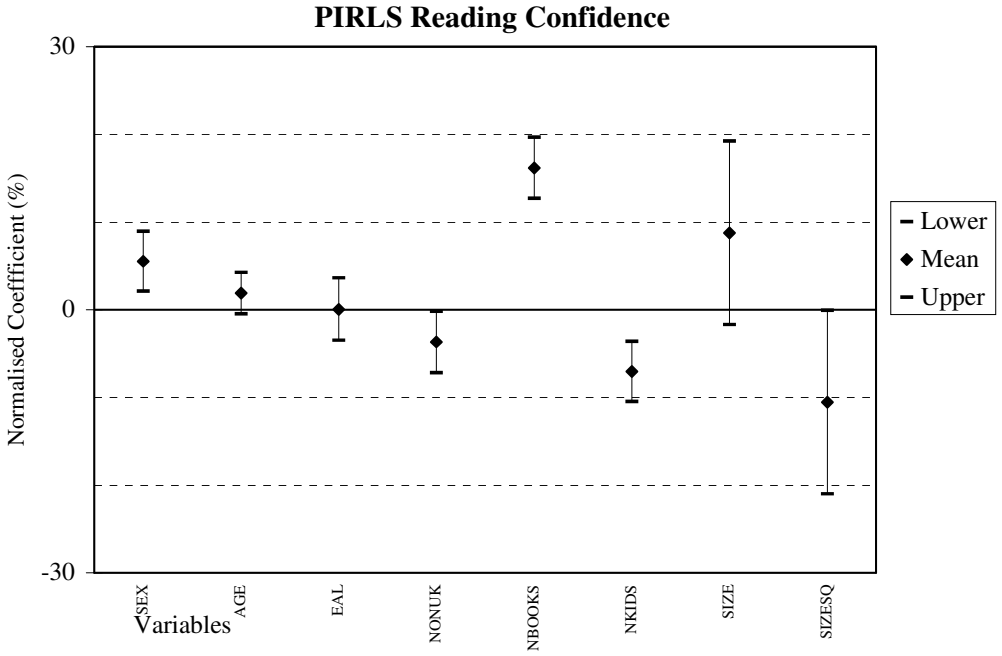


Figure 9: Normalised coefficients from Multilevel Model fitted to TV

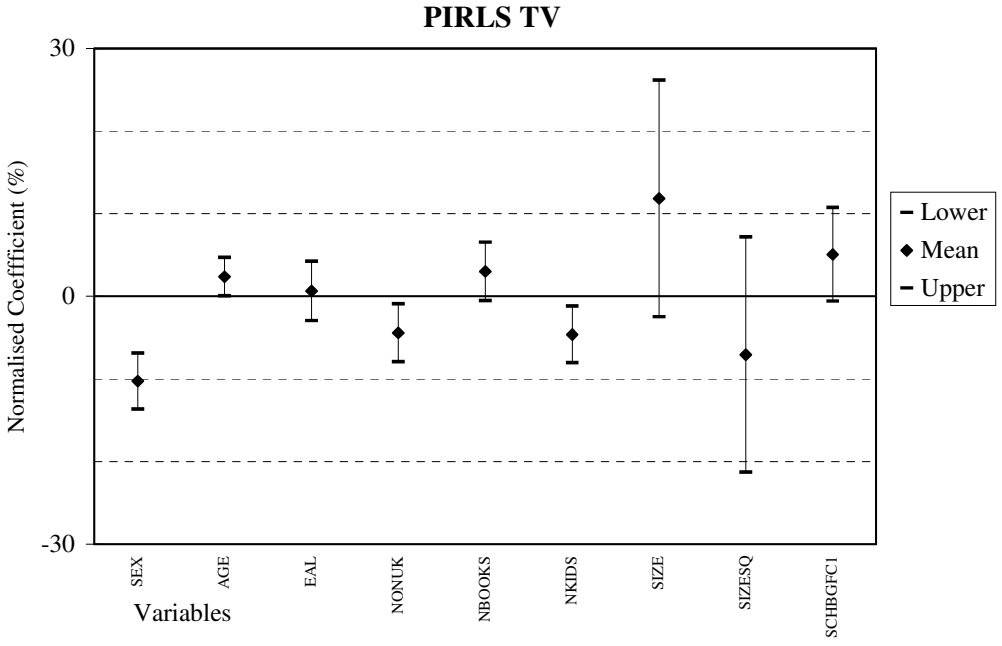
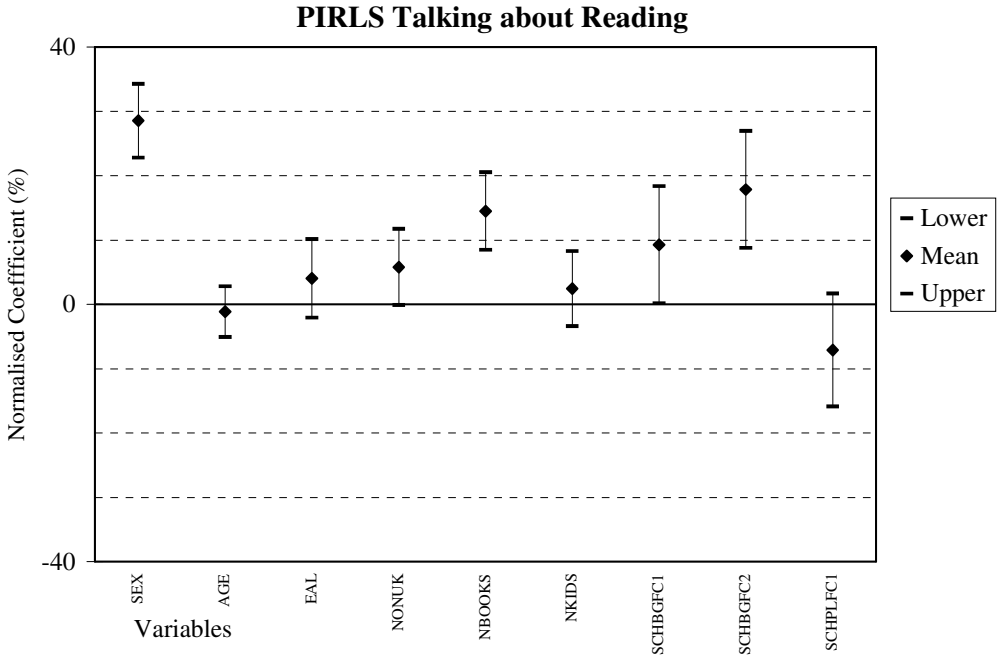


Figure 10: Normalised coefficients from Multilevel Model fitted to Talking about Reading



Effects of Pupil Variables, Excluding Attitude Factors

In the earlier modelling of pupil test scores, the analysis included both pupil background variables and attitude factors. The latter were strongly related to outcomes, so the question remains of the influence of the former when the latter are not included in the model. The analysis was rerun, for each of the three scores, with pupil background variables only included in the model. The normalised coefficients obtained are presented in Table 8 (all are statistically significant). Table 9 shows the percentage increase in the absolute value of these coefficients, compared with those derived from the full model including attitude factors (Table 6).

It is clear that the strongest relationship with test scores is for number of books in the home, followed by non-UK birth (negative). Sex and number of children in the family come next, with age and English as an additional language the lowest (but still significant).

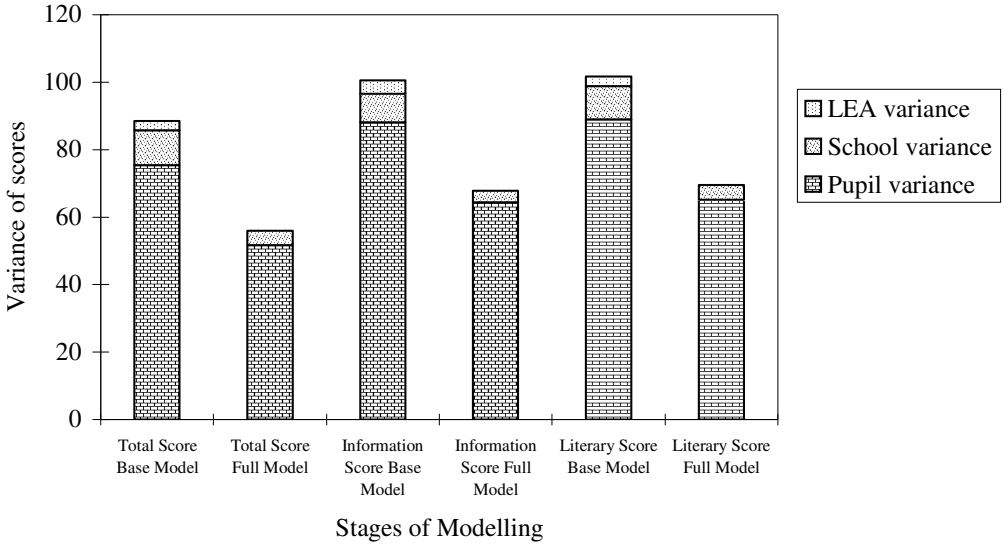
Random Variances

In addition to the relationships between international scale scores and a host of background variables described above, the multilevel models provide other information. In particular, they estimate the amount of variation in scale scores which can be attributed to different levels in the model. The amount of variation at each

level is measured by the 'variance' (basically the square of the standard deviation) at that level, and may change as extra background variables are fitted to the model.

Figures 11 illustrates this effect for the international scale outcomes. For each outcome measure, at each stage of modelling, the total variance is divided between the three levels in the model. The introduction of background variables reduces both pupil-, school- and LEA-level variances by variable amounts. For most scales, the reduction in the variation between schools is much greater than the reduction in variation between pupils, and the LEA-level variance is reduced to zero or minimal amounts.

Figure 11: Random Variances in International Cognitive Scale Scores at Different Levels, with and without Background Factors



Summary of Results for Each Scale

In this section we shall briefly summarise the findings from the multilevel analysis of each scale, in terms of the significant relationships with background factors (controlling for all other factors).

Cognitive scales (Total, Information and Literary Scores):

- Girls tended to have higher scores than boys, though for Information the difference was not significant at the 5% level;
- Older pupils tended to have higher scores on all three scales;
- Pupils born in the UK tended to have higher scores on all three scales;
- Pupils with English as an additional language tended to have lower Total and Information scores;

- Pupils in larger families tended to have lower scores on all three scales;
- Pupils with more books at home tended to have higher scores on all three scales;
- Pupils who reported higher levels of reading enjoyment and reading confidence tended to have higher scores on all three scales;
- Pupils who reported more reading activities at home and in class tended to have lower scores on all three scales;
- Pupils who reported more use of computers tended to have lower Total and Literary scores, though this was not significant for Information;
- Pupils who reported more TV watching tended to have higher scores on all three scales;
- Pupils in schools with higher percentages known to be eligible for free school meals tended to have lower scores in all three scales, though this was only significant at the 10% level for the Literary scale;
- Pupils in schools which reported a more disadvantaged background tended to have lower Literary scores, though this was only significant at the 10% level.
- 100% of the variance between LEAs, between 55% to 60% of the variance between schools , and 25% to 30% of the variance between pupils was explained by background factors.

Pupil attitude scale – reading enjoyment:

- Girls tended to have higher scores than boys;
- Pupils born in the UK tended to have higher scores;
- Pupils in larger families tended to have lower scores;
- Pupils with more books at home tended to have higher scores;
- Pupils in schools which reported higher attainment at key stage 1 tended to have higher scores.
- Pupils in schools which reported a more disadvantaged background tended to have higher scores.
- 100% of the variance between LEAs, 22% of the variance between schools, and 13% of the variance between pupils was explained by background factors.

Pupil attitude scale – reading activities in class:

- Girls tended to have higher scores than boys;
- Pupils born outside the UK tended to have higher scores;
- Pupils in larger families tended to have higher scores;
- Pupils in schools which reported higher attainment at key stage 1 tended to have higher scores, although this was only significant at the 10% level.
- Pupils in schools which reported a more disadvantaged background tended to have higher scores.
- 100% of the variance between LEAs, 32% of the variance between schools, and 3% of the variance between pupils was explained by background factors.

Pupil attitude scale – reading activities at home:

- Girls tended to have higher scores than boys;
- Pupils reporting English as an additional language tended to have higher scores, though this was only significant at the 10% level.
- Pupils born outside the UK tended to have higher scores;
- Pupils with more books at home tended to have higher scores;
- Pupils in schools which reported a more disadvantaged background tended to have higher scores.
- 26% of the variance between LEAs, 48% of the variance between schools, and 3% of the variance between pupils was explained by background factors.

Pupil attitude scale – use of computers:

- Pupils born outside the UK tended to have higher scores.
- Pupils with more books at home tended to have higher scores.
- Pupils from larger families tended to have higher scores.
- Pupils in schools with higher percentages of pupils known to be eligible for free school meals tended to have lower scores, though this was only significant at the 10% level. There was also a significant quadratic term for this variable, implying

that this effect tended to be at its lowest for a value of around 29% with a subsequent rise for higher values.

- Pupils in schools which reported a more disadvantaged background tended to have higher scores.

- None of the variance between LEAs, 23% of the variance between schools, and 2% of the variance between pupils was explained by background factors.

Pupil attitude scale – reading confidence:

- Girls tended to have higher scores than boys;

- Pupils born outside the UK tended to have lower scores.

- Pupils with more books at home tended to have higher scores.

- Pupils from larger families tended to have lower scores.

- 72% of the variance between LEAs, 3% of the variance between schools, and 4% of the variance between pupils was explained by background factors.

Pupil attitude scale – TV:

- Girls tended to have lower scores than boys;

- Older pupils tended to have higher scores.

- Pupils born outside the UK tended to have lower scores.

- Pupils with more books at home tended to have higher scores, though this was only significant at the 10% level.

- Pupils from larger families tended to have lower scores.

- Pupils in schools which reported higher attainment at key stage 1 tended to have higher scores.

- There was effectively no variance between LEAs; 10% of the variance between schools, and 2% of the variance between pupils was explained by background factors.

Pupil attitude scale – talking about reading:

- Girls tended to have higher scores than boys;

- Pupils born outside the UK tended to have higher scores.

- Pupils with more books at home tended to have higher scores.

- Pupils in schools which reported higher attainment at key stage 1 tended to have higher scores.

- Pupils in schools which reported a more disadvantaged background tended to have higher scores.

- There was effectively no variance between LEAs; 20% of the variance between schools, and 4% of the variance between pupils was explained by background factors.

Summary and Conclusions

Pupils' attainment in reading was significantly related to scales derived from their questionnaire responses, as well as to a number of background factors. The background factor with the strongest relationship to reading attainment is pupil self-reported number of books in the home; this is not a surprising finding, as it is a commonly-used surrogate for home background and cultural capital. Other factors related to reading attainment in England are non-UK birth, sex, number of children in the family, age, and having English as an additional language.

The relationships between the questionnaire-derived scales and attainment are of some interest. Reading enjoyment and confidence were positively related to attainment, as was TV watching (perhaps surprisingly). There were negative relationships between attainment and reading activities at home and in class – possibly due to poorer readers undertaking more remedial reading activities. Use of computers was also negatively related to reading attainment, though not significantly so for Information.

School level factors tended not to have significant relationships with attainment; the exception was possibly that schools reporting more disadvantaged backgrounds had lower Literary scores. School factors relating to pupil attitude scales tended also to be sparse. Those reporting a more disadvantaged background tended to be associated with higher scores for reading activities in class and at home (not surprisingly), but also for reading enjoyment (more surprisingly), use of computers (also surprisingly) and talking about reading. Schools reporting higher attainment at key stage 1 tended to be associated with higher scores for watching TV and talking about reading.

There is thus a quite complex web of relationships which has been exposed by the analysis reported here. The missing element is the prior attainment of the pupils on entry to the key stage of education – we have surrogates such as books in the home, but no direct measure. It is planned to match the national curriculum assessment results for these pupils at the end of key stage 1 (age 7) to the PIRLS data for as many as possible, and rerun some of these analyses controlling for prior attainment. This would enable us to investigate the progress made by the pupils during this period of their schooling.

It would also be useful to repeat some of these analyses, including the derivation of pupil and school scales, for other countries in the PIRLS dataset. This would enable us to see if the relationships found were confined to England only, or had more international currency.

References

GOLDSTEIN, H. (2003) **Multilevel Statistical Models (Third Edition),** London: Arnold.

TWIST, L., SAINSBURY, M., WOODTHORPE, A. & WHETTON, C. (2003) **Reading All Over the World: PIRLS National Report for England.** Slough: NFER.

Table 5: Details of Variables Used in Multilevel Modelling

Name	Min.	Max.	Description
LEA	209	938	
IDSCHOOL	1	544	*SCHOOL ID*
IDSTUD	10101	5440227	*STUDENT ID*
CONS	1	1	
TOTSCO	110	175	AVERAGE SCORE ON ALL BLOCKS
INFSCO	-99	182	AVERAGE SCORE ON INFORMATION BLOCKS
LITSCO	-99	176	AVERAGE SCORE ON LITERARY BLOCKS
SEX	1	2	Sex (0 = male, 1 = female)
AGE	9	12	Age in years at testing
EAL	0	1	English as an additional language
NONUK	0	1	Born outside UK
NBOOKS	1	5	No. of books in the home
NKIDS	1	11	No. of children at home
STMINFC1	-9	25	Student minor factor: reading enjoyment
STMINFC2	0	30	Student minor factor: reading activities
STMINFC3	0	30	Student minor factor: reading activities
STMINFC4	0	30	Student minor factor: use of computers
STMINFC5	-20	20	Student minor factor: reading confidence
STMINFC6	0	35	Student minor factor: TV
STMINFC7	-7	27	Student minor factor: talking about reading
RURAL	0	1	Rural school location
SUBURBAN	0	1	Suburban school location
SIZE	0	2	No. of Yr 5 (/100)
SIZESQ	0	4	Size squared
PCFSM	0	70	percentage eligibility for free school m
FSMSQ	0	49	% FSM squared (divided by 100)
KS198OV	1	5	KS1 Overall 1998 (5 pt scale)
SCHBGFC1	0	30	School background: Yr 1 attainment

SCHBGFC2	-11	11	School background: disadvantaged
SCHBGFC3	4	20	School background: cultural/ed. Facilities
SCHPLFC1	0	40	School policy: early introduction of reading elements
SCHPLFC2	3	25	School policy: English emphasis + parental involvement
SCHPLFC3	0	20	School policy: parents & kids influence curriculum
SCHPLFC4	-20	15	School policy: higher v. lower reading skills

Table 6: Summary of Results for International Test Scores
(Significant Normalised Coefficients)

(Significant coefficients at 5% level) **Normalised coefficients**

Variable	**Total score**	**Information score**	**Literary score**
Sex (0 = male, 1 = female)	5	*3*	8
Age in years at testing	4	4	3
English as an additional language	-4	-5	
Born outside UK	-13	-14	-11
No. of books in the home	13	10	12
No. of children at home	-4	-5	-4
Student minor factor: reading enjoyment	25	25	23
Student minor factor: reading activities	-10	-9	-10
Student minor factor: reading activities	-14	-14	-13
Student minor factor: use of computers	-3		-4
Student minor factor: reading confidence	28	26	26
Student minor factor: TV	10	10	9
Student minor factor: talking about reading			
Rural school location			
Suburban school location			
No. of Yr 5 (/100)			
Size squared			
percentage eligibility for free school m	-13	-14	*-6*
% FSM squared (divided by 100)			
KS1 Overall 1998 (5 pt scale)			
School background: Yr 1 attainment			
School background: disadvantaged			*-6*
School background: cultural/ed. Facilities			
School policy: early introduction of reading elements			
School policy: English emphasis + parental involvement			
School policy: parents & kids influence curriculum			
School policy: higher v. lower reading skills			

(Normalised coefficients in italics are significant at the 10% level)

Table 7: Summary of Results for Pupil Attitude Scales
(Significant Normalised Coefficients)

(Significant coefficients at 5% level) **Normalised coefficients**

Variable	**Reading enjoy-ment**	**Reading activities in class**	**Reading activities at home**	**Use of comput-ers**	**Reading confid-ence**	**TV**	**Talking about reading**
Sex (0 = male, 1 = female)	24	11	12		6	-10	29
Age in years at testing						2	
English as an additional language			*3*				
Born outside UK	-5	9	4	7	-4	-4	*6*
No. of books in the home	23		9	12	16	*3*	14
No. of children at home	-4	7		5	-7	-5	
Rural school location							
Suburban school location							
No. of Yr 5 (/100)							
Size squared					-11		
Percentage eligibility for free school meals				*-15*			
% FSM squared (divided by 100)			*16*	*14*			
KS1 Overall 1998 (5 pt scale)							
School background: Yr 1 attainment	6	*6*				*5*	9
School background: disadvantaged	9	24	14	8			18
School background: cultural/ed. Facilities							
School policy: early introduction of reading elements							
School policy: English emphasis + parental involvement							
School policy: parents & kids influence curriculum							
School policy: higher v. lower reading skills							

(Normalised coefficients in italics are significant at the 10% level)

Table 8: Summary of Significant Normalised Coefficients for Test Scores Related to Pupil Background Factors Only

Variable	Total score	Information score	Literary score
Sex (0 = male, 1 = female)	9	6	12
Age in years at testing	5	5	4
English as an additional language	-5	-6	-3
Born outside UK	-18	-18	-16
No. of books in the home	23	21	22
No. of children at home	-10	-10	-9

Table 9: Percentage Increase in Absolute Values of Coefficients with Pupil Background Factors Only

Variable	Total score	Information score	Literary score
Sex (0 = male, 1 = female)	74%	124%	46%
Age in years at testing	23%	18%	30%
English as an additional language	28%	28%	31%
Born outside UK	37%	29%	43%
No. of books in the home	77%	97%	77%
No. of children at home	128%	98%	135%