The Irish Journal of Education, *xxxi*, pp. 25-49.

EXPLAINING PERFORMANCE ON THE JUNIOR CERTIFICATE EXAMINATION: A MULTILEVEL APPROACH

Nick Sofroniou, Gerry Shiel and Judith Cosgrove
Educational Research Centre
St Patrick's College, Dublin

Hierarchical linear models of achievement on the Junior Certificate Examination are presented. Response variables consist of the Junior Certificate Examination grades in English, Mathematics, and Science of students who participated in the OECD Programme for International Student Assessment (PISA) in 2000, scored according to a points system. Explanatory variables consist of answers from students and the principal teachers of their schools to questionnaire items. The models compare well with earlier models of performance on the PISA reading, mathematical, and scientific literacy scales, in terms of the proportions of between-school and within-school variance they explain, and in confirming the contributions of a range of school- and student-level variables to performance, including school disadvantaged status, school type, student socioeconomic status, student gender, home educational environment, risk of dropping out of school, and attitude to reading. The results are discussed in terms of the methodological issues they raise as well as their implications for addressing underachievement among some students.

The purpose of the study reported in this paper was to develop hierarchical linear models of performance on the Junior Certificate Examination in English, Mathematics and Science, drawing on sets of explanatory variables that had been used in the development of similar models based on the performance of Irish students in reading literacy, mathematical literacy, and scientific literacy in the Organisation for Economic Cooperation and Development (OECD) Programme for International Student Assessment (PISA) (Shiel, Cosgrove, Sofroniou, & Kelly, 2001), and to compare the two sets of models. It was hypothesized that the models based on performance on the Junior Certificate Examination would have broadly similar characteristics to those developed for PISA, and hence would support the conclusions that were derived from the PISA models.

BACKGROUND

The OECD Programme of International Student Assessment

In 2000, the first cycle of an international assessment involving 15-year olds, the OECD/PISA assessment, was conducted in 28 OECD member countries (including Ireland) and in four additional countries.[1] The primary function of PISA is to generate comparative international data on students' achievements in three domains – reading literacy, mathematical literacy, and scientific literacy – and to inform the development of policy in participating countries on matters associated with achievement. Unlike earlier international studies involving school-age populations, which sought to measure students' mastery of curricular content, PISA takes a literacy-based approach that seeks to measure the cumulative yield of education at the point at which compulsory education ends in most OECD countries in terms of the knowledge and skills that students need in adult life.

In PISA 2000, the major assessment domain was reading literacy, and mathematical literacy and scientific literacy were minor domains. While the assessment of reading literacy was comprehensive, covering several important sub-processes, only limited aspects of mathematical and scientific literacy were assessed.[2] In addition to completing tests in one or more of these domains, students completed a Student Questionnaire that sought information on such variables as socioeconomic background, parents' educational attainment, and students' attitudes towards and engagement in learning. Principal teachers in participating schools completed a School Questionnaire that sought information about such variables as school management, organization of learning in schools, and resource availability and usage. In Ireland, a sample of 3,854 15-year olds in 139 schools took part. Schools were selected with a probability proportional to size in each of three strata (large, medium, and small)[3]. Then a fixed number of 15-year olds were selected at random within each school.[4] Weighted school and student response rates of 88% and 86%, respectively, were obtained, after replacement.

1 One OECD member country, The Netherlands, was not included in several analyses in the PISA 2000 international report because the school-level response rate in that country fell below predefined international standards. The four non-OECD countries that participated in PISA 2000 are Liechtenstein, the Russian Federation, Latvia, and Brazil.

2 In 2003, mathematical literacy will be a major PISA domain, while reading literacy and scientific literacy will be minor domains. In 2006, scientific literacy will assume the status of a major domain, while reading literacy and scientific literacy will become minor domains.

3 Small schools were defined as those with 17-40 15-year olds, medium as those with 41-80, and large as those with more than 80.

4 35 students (or the actual number of students, if lower) were selected in each school.

The PISA 2000 International Report

The international report on PISA 2000 (OECD, 2001) provides a broad overview of the outcomes of the assessment, together with hierarchical linear models of achievement on the three literacy scales. A three-level model for reading literacy is presented that explains 43.4% of between-country variation, 71.9% of between-school variation, and 12.4% of the variation between students. However, the analysis omits a large number of variables that might be relevant to understanding reading achievement in individual countries. In addition, the fitting of a Normally-distributed random effects model at the country level raises theoretical issues concerning inferences about the population from which the countries are deemed to be a sample (in a frequentist statistical framework), or about the exchangeability of one country for another (in a Bayesian statistical framework) (see Raudenbush, Cheong, & Fotiu, 1994). Using the subset of international data for the Irish sample, the study described in the present paper includes a much wider range of explanatory variables, some of which were locally collected (e.g., school designated disadvantaged status), and so allows the development of models that inform explanations of the variation amongst students and schools in the Irish context.

The PISA 2000 National Report for Ireland

In the Irish national report on PISA 2000 (Shiel et al., 2001), over 40 variables judged to be relevant to students' achievement were identified and described, and associations between the variables and student achievement were reported based on bivariate statistics. Variables at the student level included student gender, student socioeconomic status, parental educational attainment, home learning environment, and student learning habits and attitudes. Variables at the school level included gender composition of the school (all males, all females, mixed), disadvantaged status, autonomy in decision making, school learning climate, and availability of resources.

Following a decision to develop hierarchical linear models of achievement in reading, mathematical, and scientific literacy, subsets of variables were identified as candidates for inclusion in the model. These were selected with reference to collinearity among variables in the full set, and policy issues identified by the National Advisory Committee for PISA. All three models (reading literacy, mathematical literacy, scientific literacy) explained over 74% of between-school variance. The model for reading literacy explained 44.2% of within-school variance, while those for mathematical and scientific literacy explained 31.9% and 34.1% respectively.

The model for reading literacy included school-level variables (disciplinary climate, school type, and disadvantaged status), student-level variables (gender, socioeconomic status, number of siblings, index of books in the home, dropout risk, frequency of absence from school, completion of homework on time, current grade level, frequency of leisure reading, and a variable reflecting the interaction between gender and number of books in the home). The model confirmed the associations of a number of variables with achievement, and indicated their estimated contributions to the fitted values for students' scores. For example, attendance at a vocational rather than a community/comprehensive school was estimated to contribute –20.4 points, or minus one-fifth of a standard deviation to a student's score. Similarly, attendance at a school designated as disadvantaged was estimated to contribute –22.3 points (about one-quarter of a standard deviation), while dropout risk was estimated to contribute –54.4 points (over one-half of a standard deviation). Variables with positive contributions included socioeconomic status, with the average contribution ranging from +25.9 points (one-quarter of a standard deviation) for students at the mean of the high SES category (the top third of the distribution of SES scores) to +3.0 points for students at the mean of the low SES (those in the bottom third). Since the model was additive, it was possible to estimate the contributions of combinations of variables.

The hierarchical models for mathematical literacy and for scientific literacy were less complex than for reading literacy, reflecting their smaller sample sizes and fewer candidate variables. School type and disadvantaged status were the only school-level variables that remained in the final models. However, these, together with student variables, accounted for sizeable proportions of between-school variance. Both models also included parents' combined socioeconomic status, an index of books in the home, grade level, completion of homework on time, and dropout risk. In the model of scientific literacy, a binary variable describing whether or not the student studied science at school contributed +43.1 points (about one-half of a standard deviation) to the scores of students who studied the subject.

The Junior Certificate Examination

In the Irish national report on PISA 2000, some comparisons were drawn between the performance of students on the tests of reading literacy, mathematical literacy and scientific literacy, and their performance on the corresponding Junior Certificate Examination subjects (English, Mathematics and Science). To establish these comparisons, the Junior Certificate Examination grades of PISA students who sat for the examination in 1999

(33.1% of the PISA cohort of 15-year olds) or 2000 (60.9%) were placed on a Junior Certificate Performance Scale (JCPS) that ranges from 1-12, and has a three-grade overlap between examination levels (see Table 1). For example, the higher-level grades, D, E, and F are deemed to be worth 9, 8, and 7 respectively, and to be equivalent to ordinary level grades A, B, and C. The scale has been used in previous research (Kellaghan & Dwan, 1995; Martin & Hickey, 1993), though in the context of describing overall performance on the examination rather than performance on individual subjects. A comparison of the 12-point scale with alternative 10- and 14-point scales for each subject area (English, Mathematics, and Science) using OLS regression revealed that the 12-point scale generally worked better than either the 10- or 14-point scales (Shiel et al., 2001, Table A6.2). It should be noted, however, that the 12-point scale is bounded at its upper and lower points. Thus, some care is needed in predicting from the extremes of explanatory variables with linear models.

TABLE 1

JUNIOR CERTIFICATE OVERALL PERFORMANCE SCALE (OPS)

Higher Level	Ordinary Level	Foundation Level	OPS Score
A			12
B			11
C			10
D	A		9
E	B		8
F	C		7
	D	A	6
	E	B	5
	F	C	4
		D	3
		E	2
		F	1

Correlations between students' scores on the PISA tests and their performance scores on the corresponding Junior Certificate subjects were .74 (reading literacy/English), .73 (mathematical literacy/mathematics) and .73 (scientific literacy/science). All of these correlations are significant ($p<.001$).[5]

5 The method for establishing the significance of correlation coefficients, along with the test statistics, is reported in Section A4.6 (p. 206) in Shiel et al. (2001).

Between-School and Within-School Variance in Achievement

One approach to examining the extent to which achievement varies with the school that students attend is to calculate the proportion of variance in achievement scores that can be attributed to schools (between-school variance) and the proportion that can be attributed to individuals within schools (within-school variance). A statistic we term the intra-cluster correlation (ICC) describes the total variation in achievement between clusters (schools, in this instance) as a percentage of the total variation within a country. According to Postlethwaite (1995), when the ICC (i.e., between-school variance) is high, 'the difference between schools is high, and, in this case, the level of a student's score will very much depend on which school he/she attends. However, when [between-school variance] is low, it does not matter what school a student attends' (p. 81). Postlethwaite reported an ICC of .48 for Irish 14-year olds in the IEA Reading Literacy Study (IEA/RLS), indicating that 48% of the variance in the their scores could be attributed to differences between schools, while the remainder could be attributed to variation within schools. However, since intact classes were sampled in IEA/RLS, it is likely that the ICC includes both between-school and between-class variance.

In an analysis of data for Irish students in eighth grade (second year in secondary school) in the Third International Mathematics and Science Study (TIMSS), carried out at the Educational Research Centre using maximum likelihood estimation of variance components,[6] ICCs of 0.44 for Mathematics and 0.33 for Science were obtained.[7] When performance on the Junior Certificate Examination in mathematics and science for the same students was analysed, ICCs of 0.61 and 0.52 were obtained. Again, however, these estimates are probably inflated by the sampling of intact classes in TIMSS.

Using overall performance on the Junior Certificate Examination as the response variable, Smyth (1999) reported an intra-school correlation of 0.22. In Smyth's study, students had been selected with reference to their base classes, and would not have been together for most subjects.

In an earlier study, Kellaghan et al. (1979) reported on between-school, between-class, and within-class variance components for a sample of 1,560 boys

6 All five plausible values were used in the TIMSS variance component calculations.

7 These are somewhat lower than the intra-cluster correlations reported for Irish students in eighth grade mathematics and science by the TIMSS Study Centre which were .50 and .38 respectively (Martin et al., 2000).

in 47 schools on individual subjects in the Intermediate Certificate Examination.[8] Estimates of the fraction of variance attributed to schools ranged from .28 (Higher-level English) to .61 (Higher-level Irish), with an estimate of .32 for overall performance on the examination. Between-school estimates were generally higher than between-class estimates. The relatively high estimates of variance associated with schools were attributed to the selective nature of second-level schools.

Finally, in the PISA 2000 study, in which 15-year olds were selected at random within schools, the intra-cluster (school) correlations were .18 for reading literacy, .11 for mathematical literacy, and .14 for scientific literacy (OECD, 2001, Table 2.4, Table 3.5). The corresponding OECD country averages were .35, .31 and .31 respectively. Just four OECD countries (Iceland, Norway, Sweden and Finland) had lower intra-cluster correlations than Ireland for mathematical and scientific literacy.

METHOD

Subjects

Subjects in the current study were 15-year olds in Irish post-primary schools who participated in OECD/PISA in 2000, who sat the Junior Certificate Examination (JCE) in either 1999 or 2000, and for whom data on all the PISA student-level candidate variables (see Table 2 below) were available.[9] In all, scores for 3,428 students were available for Junior Certificate English, 1,896 for Junior Certificate Mathematics, and 1,629 for Junior Certificate Science. The proportions of male students in the English, Mathematics and Science analyses were .47, .47, and .51 respectively.

The mean Junior Certificate Performance Scale score for English was 9.25 (N = 3,428, SD = 1.65); the mean score for Mathematics, 8.20 (N = 1,896, SD = 2.05); and the mean score for Science, 9.36 (N = 1,629, SD = 1.77).

8 The Intermediate Certificate Examination was subsequently replaced by the Junior Certificate Examination.

9 A listwise deletion procedure was selected in modelling using the HLM package. The procedure results in the omission of any students with missing values on any of the selected student-level variables. The total numbers of students for whom JCE scores were available were 3,625 for English, 2,002 for Mathematics and 1,720 for Science. Hence, the listwise deletion resulted in the loss of about 5% of cases for each subject area.

Selection of Candidate Variables

Candidate variables considered for modelling students' achievement in Junior Certificate English, Mathematics, and Science were ones which had separately shown significant associations with the PISA 2000 achievement scales, with one exception (see below). The variables are listed in Table 2, and are described in greater detail in Appendix 1.

TABLE 2

LIST OF CANDIDATE STUDENT AND SCHOOL VARIABLES

Student Variables	School Variables
Background	School Structure
Gender	School Size (Stratum)
Socioeconomic Status (Combined Scale)	School Type
Parents' Education (Combined Scale)	Designated Disadvantaged Status
Lone Parent Status	Gender Composition
Number of Siblings	
Home Educational Climate	School Climate/Policy
Parental Engagement	Negative Disciplinary Climate
Index of Number of Books in the Home	
Student as Learner	School Resources
Dropout Risk	Student-teacher Ratio
Absence from School	
Frequency of Homework Completed	
Junior Cert Year (1999 or 2000)	
Reading Habits and Attitudes (included in model for English only)	
Diversity of Reading	
Frequency of Leisure Reading	
Attitude towards Reading	

Some of the variables are composites that were constructed for each OECD country, and represent a number of individual variables which are closely associated. For example, the variable, 'parental engagement' was constructed by combining responses to a number of statements regarding the frequency with which parents engaged with students in various activities, such as discussing politics, books, films, and television programmes. Where several variables in the PISA database were correlated or closely linked theoretically (for example, mother's socioeconomic status, father's socioeconomic status, parents' combined socioeconomic status), a single variable was chosen as a candidate variable for our models, either because it had the strongest correlation with achievement, or because it was of particular policy interest.

While most variables were derived from the School and Student Questionnaires administered in PISA, a few that were of particular interest in the Irish context, such as school type (whether secondary, community/ comprehensive, or vocational), and school designated disadvantaged status (whether or not a school is designated as disadvantaged by the Department or Education and Science), were derived from the Department of Education and Scienc database of post-primary schools.

One variable used in modelling achievement on PISA, but not used in the current study, was a student's current grade level. This was not used because it is very highly correlated with another variable, Junior Certificate Year – whether a student took the Junior Certificate Examination in 1999 or 2000. Three variables – diversity of reading, frequency of leisure reading, and attitude towards reading – were considered for inclusion in the model for English only.

Implementation of Modelling Procedures

Models for each subject (English, Mathematics, Science) were developed separately using hierarchical multilevel modelling. Such models incorporate an additional random component at the level of the cluster (e.g., schools in PISA) that allows for the variation present across clusters to be taken into account. In addition, one or more random coefficients can be included in the models. A random component at the school level consisting of just a random intercept indicates the slopes of the fitted parameters are constant, but that they vary in a parallel manner from school to school. Adding a random slope for a term in the model indicates that the slope for that explanatory variable also varies across schools.

In the models presented here, full maximum likelihood estimation was used, enabling deviance tests of both fixed and random effects to be carried out. In line with the advice of Aitkin et al. (in press), sampling weights were not applied in developing the models, and the design strata were incorporated in the model building process. As in the models reported in the Irish national PISA report, uncentred continuous variables were used. Hence, the intercept has the conventional interpretation of ordinary least-squares regression (i.e., the value of the linear predictor when the continuous explanatory variables are set to zero). The development of each model involved the following procedures.

First, candidate variables at the student level were evaluated separately as fixed effects added to a random intercept-only model of overall achievement in English on the Junior Certificate Examination, and then evaluated jointly. Non-significant variables (with the exception of gender, to enable later evaluation of any gender interactions which are of policy interest) were omitted manually using a backwards elimination strategy. Any borderline variables ($p \leq .1$) were retained in the model at this stage. Categorical variables with more than two levels were evaluated by omnibus tests of deviance changes, fitting the model with and without the corresponding set of dummy variables. Next, interactions between gender and each of the level 1 (student) variables were tested separately by addition to the model and significant interactions were added to form a new model.

Then level 2 (school) variables were tested separately by adding each one to a random intercept-only model as a fixed effect, and eliminating those that were not significant.[10] The explicit stratifying variable school size (the number of 15-year olds), which had been used in the sampling design, was included as a level 2 categorical variable, as were two implicit stratifying variables, school gender composition and school type. The evaluation of these variables served to ensure that the unweighted analysis was not distorted by over-sampling of any particular subgroup. The level 2 variables were added to the random intercept model containing the remaining level 1 variables. Variables for which parameter estimates were not statistically significant were removed sequentially. A stricter criterion ($p \leq .05$) was applied for retention of the remaining explanatory variables.

Following this, estimates and tests of quadratic terms for level 1 and level 2 continuous variables were tested separately, or jointly in cases where there were interactions between continuous variables and gender. Significant quadratic terms and their interactions with gender were added to the model and evaluated sequentially. Finally, random coefficients for level 1 variables were tested through deviance changes referred to a chi-squared distribution, with the number of degrees of freedom corresponding to the change in the number of terms in the model. Significant random effects were added to the model.

10 Since one of the level 2 variables (student-teacher ratio) had missing data for one school or 1% of cases in each examination subject, a missing value indicator method was applied in which student-teacher ratio was nested within a binary variable where a value of 1 indicated non-missing. The purpose of this was to prevent the loss of students from schools for which the variable student-teacher ratio was not available.

RESULTS

The development of the model of performance on Junior Certificate Examination English is presented in detail first and estimates of the effects of example values of variables used in the final model are given to facilitate interpretation of results. Then the outcomes for the JCE Mathematics and JCE Science models are presented briefly.

Model of Performance on Junior Certificate Examination English

Prior to testing each level 1 variable separately by addition to the random intercept only model, a preliminary investigation of the curvilinearity of their relationships with the JCE English performance scores suggested that one variable, index of the number of books in the home, would be well represented by its logarithmic form. Table 3 gives the parameter estimates and tests of statistical significance of the level 1 variables that were initially tested separately.

When all the variables represented in Table 3 were simultaneously entered into the same random intercept model, the variable diversity of reading (which reflects the range of materials that students read) was found not to be significant and was removed. Two of the three parameter estimates for frequency of reading (30-60 minutes per day, more than 60 minutes) were then observed to have changed sign from positive to negative. As occurred when modelling PISA combined reading literacy, these two parameter estimates are positive when fitted alone or together with each of the other variables. However, their signs are negative when the variables attitude to reading and frequency of reading are included in the same model.

After testing for gender interactions with each of the remaining level 1 variables, two interactions were added to the model: gender by attitude to reading, and gender by the index of the number of books in the home.

The next phase involved evaluating each of the level 2 variables separately when added to the intercept-only model (see Table 4), with student-teacher ratio being tested as a fixed effect nested within its non-missing indicator variable.

In the next model, which included the additional level 2 variables in Table 4, the categorical variable school size (number of 15-year olds in the school), which was an explicit sampling variable, was not significant. In the following model, the deviance test for the level 2 categorical variable school gender composition indicated non-significance, suggesting that it should also be eliminated.

TABLE 3

ACHIEVEMENT ON JUNIOR CERTIFICATE EXAMINATION ENGLISH (PERFORMANCE SCORES): ALL LEVEL 1 VARIABLES TESTED AS SEPARATE MODELS BY ADDITION TO THE RANDOM INTERCEPT ONLY MODEL

	Parameter	*SE*	*Test Statistic*	*df*	*p*
Gender: Male – Female	–0.758	0.068	t = –11.075	3426	<. 001
Socioeconomic Status	0.025	0.002	t = 14.451	3426	<. 001
Parental Education			Ddiff =113.858	3	<. 001
None/Primary–Upper Sec	–0.449	0.093			
Lower Sec–Upper Sec	–0.350	0.073			
Third Level–Upper Sec	0.325	0.064			
Lone Parent: Yes– No	–0.337	0.082	t =–4.089	3426	<. 001
Number of Siblings	–0.119	0.019	t =–6.253	3426	<. 001
Parent Engagement	0.327	0.027	t =11.918	3426	<. 001
Log (Index of Books in the Home)	1.383	0.078	t =17.700	3426	<. 001
Dropout Risk: Yes– No	–1.660	0.077	t =–21.508	3426	<. 001
Absence			Ddiff = 70.189	2	< 001
No days–1 or 2 days	0.302	0.057			
Three days or more–1 or 2 days	–0.428	0.100			
Homework on Time			Ddiff =164.051	3	<. 001
Never–Mostly	–0.836	0.126			
Sometimes–Mostly	–0.566	0.065			
Always–Mostly	0.292	0.066			
Junior Cert. Year (2000-1999)	–0.307	0.056	t =–5.497	3426	<. 001
Diversity of Reading	0.347	0.056	t =–5.497	3426	<.001
Freq. Of Leisure Reading			Ddiff =183.093	3	<.001
No time–Up to 30 mins	0.644	0.064			
30–60 mins–Up to 30 mins	0.189	0.072			
>60 mins–Up to 30 mins	0.162	0.080			
Attitude to Reading	0.546	0.025	t =21.619	3426	<.001

In two subsequent models, the variables student-teacher ratio (nested within its non-missing indicator variable), and its non-missing indicator variable were both non-significant and therefore were removed. The remaining variables were all significant. At this point, estimates of the quadratic terms for each of the remaining level 1 and level 2 variables were tested separately, except interaction terms related to attitude to reading, which were tested jointly. The quadratic terms for number of siblings, attitude to reading, and gender by attitude to reading were then added to the model. Finally, random coefficients for the remaining level 1 variables were tested by adding each one separately to the model, yielding a final model (see Table 5), which includes two random coefficients at the student level (Junior Certificate year and dropout risk). The variance associated with each of these random variables is statistically significant when tested by removal from the model.

TABLE 4

ACHIEVEMENT ON JUNIOR CERTIFICATE EXAMINATION ENGLISH (PERFORMANCE SCORES): ALL LEVEL 2 VARIABLES TESTED AS SEPARATE MODELS BY ADDITION TO THE INTERCEPT-ONLY MODEL

	Parameter	*SE*	*Test Statistic*	*df*	*p*
Negative Disciplinary Climate	–0.637	0.165	t = –3.872	137	< .001
School Type			Ddiff = 50.863	2	< .001
Secondary–Community/Comp	0.424	0.162			
Vocational–Community/Comp	–0.602	0.186			
Not Designated Disadv–Disadvantaged	0.886	0.128	t = 6.911	137	< .001
School Gender Composition			Ddiff = 22.073	2	< .001
All Males–Mixed	0.214	0.169			
All Females–Mixed	0.697	0.142			
School Size (Number of 15-Year Olds)			Ddiff = 7.840	2	0.020
Large–Medium	0.020	0.302			
Small–Medium	0.401	0.151			
Student-Teacher Ratio	0.110	0.033	t = 3.337	136	< .001

The intra-cluster correlation (the proportion of total variability due to the cluster, i.e., school level) is 0.177. This indicates that 17.7% of the variance in achievement can be attributed to the school level, and that the remainder can be attributed to the student/class levels, though no information on the assignment of students to classes was gathered in PISA. To estimate the proportion of variance in student achievement explained by the model at both school and student levels, the variance components associated with the model prior to the addition of the random coefficients for Junior Certificate year and dropout risk were used.

TABLE 5

FINAL MODEL OF ACHIEVEMENT ON JUNIOR CERTIFICATE EXAMINATION ENGLISH (PERFORMANCE SCORES)

	Parameter	*SE*	*Test Statistic*	*df*	*p*
Intercept	7.588	0.215			
Student-Level Variables					
Gender: Male–Female	0.112	0.228			
Socioeconomic Status	0.013	0.002	t = 8.120	3401	<.001
Lone-Parent Status: Yes–No	–0.151	0.070	t = –2.145	3401	0.032
Log (Index of Books in the Home)	0.818	0.106			
Number of Siblings	0.110	0.051			
Number of Siblings Squared	–0.032	0.008	t = –3.969	3401	<.001
Parental Engagement	0.079	0.025	t = 3.135	3401	0.002
Dropout Risk: Yes–No	–1.071	0.085	t = –12.592	3401	<.001
Junior Certificate Year (2000-1999)	–0.209	0.055	t = – 3.789	3401	<.001
Absence			Ddiff = 11.902	2	0.003
No days – 1 or 2 days	0.093	0.049			
Three days or more – 1 or 2 days	–0.182	0.086			
Homework on Time			Ddiff = 28.473	3	0.013
Never–Mostly	–0.190	0.112			
Sometimes–Mostly	–0.235	0.057			
Always–Mostly	0.116	0.057			
Freq. Of Leisure Reading			Ddiff =33.804	3	<.001
No time–Up to 30 mins	–0.010	0.064			
30-60 mins–Up to 30 mins	–0.164	0.065			
> 60 mins–Up to 30 mins	–0.456	0.079			
Attitude to Reading	0.425	0.043			
Attitude to Reading Squared	0.004	0.020			
Gender X Attitude + (Gender X Attitude) Squared			Ddiff =7.362	2	0.025
Gender X Attitude	–0.099	0.052			
(Gender X Attitude) Squared	0.060	0.030			
Gender X Log (Index of Books in the Home)	–0.350	0.142	t = – 2.463	3401	0.014

TABLE 5 – CONTINUED

	Parameter	*SE*	*Test Statistic*	*df*	*p*
School-Level Variables					
Negative Disciplinary Climate	–0.188	0.091	t =– 2.062	134	0.039
School Type			Ddiff =16.385	2	<.001
Secondary-Community/Comp	0.004	0.102			
Vocational-Community/Comp	–0.350	0.115			
Not Designated Disadv–Disadvantaged	0.499	0.080	t = 6.239	134	<.001
Variance Components					
Level 2 Random Component (RC)					
Intercept Variance	0.083				
Dropout Risk RC			Ddiff = 7.887	3	0.048
Dropout Risk Variance	0.195				
Dropout Risk-Intercept Covariance	–0.073				
Junior Certificate Year RC			Ddiff =13.907	3	0.003
Junior Cert Variance	0.107				
Junior Cert Intercept Covariance	–0.028				
Junior Cert – Dropout Risk Covariance	0.072				
Level 1 Variance	1.605				

Variables Dropped from Model (in sequence)	
Diversity of Reading	Gender Composition of School
Parental Education	Student-Teacher Ratio x Non-Missing
School Size	Non-Missing Indicator for Student-Teacher Ratio

Note: In the above model, the interactions of gender with the linear and quadratic terms for attitude to reading were tested in a single omnibus test. The value of the gender by attitude to reading squared term was also tested individually and was found to be statistically significant (t = 1.973; df = 3401; p = 0.048).

Using the mean number of 15-year olds enrolled in the population (86.88) as a representative cluster size, level 1 R^2 is 0.792, while level 2 R^2 is 0.373.[11] Hence,

11 See Snijders & Bosker (1999) for an explanation of the calculation of multilevel R^2 statistics.

the final model explained 79.2% of the variance in achievement at the school level, and 37.3% at the student level. It was estimated that the additional variance explained by the addition of level 2 variables to the model containing only student-level variables was 17.3% at the school level, and 3.2% at the student level.

Contribution to Fitted Values of the Model for Junior Certificate Examination English

The contribution of a number of variables to fitted values from the final model of performance in Junior Certificate Examination English is examined in this section. Example values are useful when the individual parameter estimates cannot be translated directly into units of the response variable (i.e., quadratic fits and variables involved in interactions). Example values for continuous variables are also given (for example, negative disciplinary climate), even though their parameters do have a direct interpretation.

The model is additive in the sense that every variable makes an added contribution to the linear predictor. The contributions of categorical variables to the linear predictor are immediately apparent from the final model (Table 5). A student in a vocational school has a predicted JCE English grade that is one-third of a grade (one-fifth of a standard deviation) lower than that of a student in a community/comprehensive school. A student in a school that is not designated as disadvantaged is likely to score one half of a grade (just under one-third of a standard deviation) higher than a student in a school designated as disadvantaged. A student at risk of dropping out of school is expected to score about one grade point lower in English (almost two-thirds of a standard deviation) than a student who is not at risk. A student who sat the JCE English examination in 2000 is expected to score one-fifth of a grade (just over one tenth of a standard deviation) lower than a student who sat the examination in 1999.

Continuous variables were categorized into high, medium, and low categories, using the values corresponding to the 33rd and 67th percentiles on their scales as cut-points, and these values were used in estimating effects of being at the mean of each of these groups, using the parameter values from the model in Table 5.[12] Table 6 gives the contributions for students scoring at the means of the high, meduim, and low categories for socioeconomic status. Scoring at the mean of the 'low' category is associated with an increment of over one-third of a grade (almost one-quarter of a standard deviation). The

12 The methodology for this is described in Inset 5.4 (p.107) in Shiel et al. (2001).

corresponding estimates for the medium and high groups are three-fifths of a grade (over one-third of a standard deviation) and four-fifths of a grade (almost one-half of a standard deviation) respectively.

TABLE 6

CONTRIBUTIONS TO GRADES IN JUNIOR CERTIFICATE EXAMINATION ENGLISH ATTRIBUTABLE TO SOCIOECONOMIC STATUS

Level of SES	Estimated Contributions to Scores
Low SES	0.386
Medium SES	0.591
High SES	0.814

In the case of parental engagement, a variable that reflects the frequency with which parents engage with students in discussing politics and social issues, discussing books, films or television programmes, and listening to classical music, the effects on students' grades in JCE English for students with low, medium, and high levels of engagement are all less than one-tenth of a grade (one-twentieth of a standard deviation) (Table 7). Although these effects are statistically significant, they are very small in comparison with those for other variables in the model.

TABLE 7

CONTRIBUTIONS TO GRADES IN JUNIOR CERTIFICATE EXAMINATION ENGLISH ATTRIBUTABLE TO PARENTAL ENGAGEMENT

Parental Engagement	Estimated Contributions to Scores
Low	–0.096
Medium	0.009
High	0.063

Tables 8 and 9 show the contributions to the linear predictor of students' grades in JCE English for attitude to reading (which has a squared term, as well as an interaction with gender) and the index of books in the home (which is in its logarithmic form, and also has an interaction with gender). The estimates in Table 8 indicate that poor (negative) attitude to reading appears to be less detrimental to a male student (lowering his performance by less than one-fifth of a grade or one-tenth of a standard deviation) than to a female student (whose performance drops by about one-half of a grade or over one-quarter of a standard deviation). Table 9 suggests that the index of books in the home (a proxy for the quality of the home educational environment) has a somewhat stronger effect on female students than on male students. For example, the estimated contribution for male students with between 251 and 500 books is almost one grade (over one-half of a standard deviation), while the corresponding effect for female students is just under one and a half grades (nine-tenths of a standard deviation). Similarly, the effect of the index of books in the home reaches an asymptote more rapidly for male students than for female students.

TABLE 8

CONTRIBUTIONS TO GRADES IN JUNIOR CERTIFICATE EXAMINATION ENGLISH ATTRIBUTABLE TO ATTITUDE TO READING, FOR MALE AND FEMALE STUDENTS

Attitude to Reading	Estimated Contributions to Scores	
	Male	Female
Poor Attitude	–0.162	–0.447
Average Attitude	0.089	–0.031
Good Attitude	0.529	0.455

TABLE 9

CONTRIBUTIONS TO GRADES IN JUNIOR CERTIFICATE EXAMINATION ENGLISH ATTRIBUTABLE TO NUMBER OF BOOKS IN THE HOME, FOR MALE AND FEMALE STUDENTS

Index of Books in the Home	Estimated Contributions to Scores	
	Male	Female
No Books (1)	0.112	0.00
1-10 books (2)	0.437	0.567
11-50 books (3)	0.627	0.899
51-100 books (4)	0.761	1.134
101-250 books (5)	0.866	1.317
251-500 books (6)	0.952	1.466
+500 books (8)	1.024	1.592

For a variable with a random slope, the range of values that the slope takes for 95% of the population may be estimated by taking the square root of the variance associated with the slope (i.e., its standard deviation) and adding +/–1.96 times this value to the parameter estimate. In the case of dropout risk, the range is –1.94 to –0.21. It would therefore be expected that, in 95% of schools, being in the group at risk of dropping out would result in grades that are between one-tenth and one and one-fifth of a standard deviation lower than the grades of students in the group that are not at risk. In the case of Junior Certificate year, the values that the slope takes are in the range of –0.85 to +0.43. This indicates that, in some schools, being in the group taking JCE English in 2000 would result in achievement of a lower JCE English grade than students in the group that took the examination in 1999, while in other schools, there would be an increase in performance. An average school would have an expected difference between 2000 and 1999 equal to the parameter estimate of –0.209.

Models of Performance on Junior Certificate Examination Mathematics and Science

The final models for Junior Certificate Examination Mathematics and Science are broadly similar to the final model for English. The intra-cluster correlation for JCE Mathematics is 0.156, while that for JCE Science is 0.162. The final model of performance in JC Mathematics explains 64.1% of the variance in achievement at the school level, and 29.5% at the student level. The final model for JCE Science explains 71.3% of variance at the school level, and 31.1% at the student level. The school-level variables in the final model for JCE Science are the same as those for JCE English. An additional variable, negative disciplinary climate, was retained for JCE Mathematics.

The final model for JCE Mathematics includes an interaction for gender by completion of homework on time. Whereas female students who never completed homework on time tended to do worse than the corresponding males (by four-fifths of a grade level, or two-fifths of a standard deviation), female students who always completed their homework on time tended to do better (by about one-tenth of a grade) than males who always completed their homework on time. The final model for JCE Science includes an interaction between gender and the log of the index of books in the home. The effects for male students are larger than for females, up to and including 10-50 books. For higher numbers of books, the effects for female students are greater. In English, this cross-over occurred at a lower point on the index, i.e., 1-10 books. Whereas the variable parental education (combined scale) had been dropped from the model for JCE English, it was retained in the models for JCE Mathematics and JCE Science.

CONCLUSION

The intra-cluster (school) correlations in PISA are .178 for reading literacy, .114 for mathematical literacy, and .141 for scientific literacy. These are below the corresponding OECD country means, suggesting that, compared with schools in other countries, Irish schools are relatively homogeneous in terms of achievement, but that there is considerable variation in achievement within schools. The intra-cluster correlations for the Junior Certificate Examination are 0.177 for English, 0.156 for Mathematics, and 0.162 for Science. Hence, the proportions of between-school variance for the three Junior Certificate Examination subjects are similar to those for the three PISA literacy domains (despite, for example, differences between PISA reading literacy and Junior Certificate English), and are considerably lower than those reported for earlier studies such as the IEA Reading Literacy Study and the Third International Mathematics and Science Study, where different approaches to sampling had been used (intact classes were selected).

In the current study, the final model for JCE English explains 79.2% of between-school variance and 37.3% of within-school variance. The corresponding estimates for JCE Mathematics are 64.3% and 29.5%, while those for JCE Science are 71.3% and 31.1%. Hence, for the most part, the JCE models explain similar proportions of between-school and within-school variance to the PISA models for reading, mathematical, and scientific literacy.

The final models of performance on PISA reading literacy and JCE English are broadly similar. The effect for non-designated schools is one-quarter of a standard deviation in PISA and just under one-third of a standard deviation (one-half of a grade) in JCE English. In both models, relatively small positive effects are apparent for secondary schools over community/comprehensive schools, but larger negative effects are observed for vocational schools – one-fifth of a standard deviation on both PISA and JCE English scales. These findings underline the importance of certain school characteristics on achievement. In particular, it seems that a student who attends a vocational school that is designated as disadvantaged is particularly at risk of low achievement. Interestingly, school gender composition is not included in either of the final models, as its effects are accounted for by other variables in the models.

While student socioeconomic status has a significant positive effect in both the PISA reading literacy and JCE English, other home-related variables, such as the index of the number of books, number of siblings, and attitude to reading also have significant effects. The inclusion of both attitude to reading and frequency of reading in the models has the effect of rendering the contributions of high frequencies of reading for leisure negative. This is perhaps counter-intuitive, in that a positive attitude to reading and high frequency of leisure reading might both be expected to make positive contributions to performance. However, it may be the case that, once attitude to reading is taken into account, high amounts of leisure reading are associated with lower performance on assessments involving reading. Certainly, some 15-year olds, who are quite good readers, may exercise their reading skills in content area reading texts, while others, with somewhat lower achievement, may spend larger amounts of time engaged in leisure reading. Agresti and Finlay (1997) have referred to the phenomenon of a change in sign of the association between an explanatory variable and a response variable with the addition of a second variable (attitude to reading in this instance) as Simpson's paradox. In any event, this relationship needs to be examined further.

There is an interaction between gender and the logarithm of the index of books in the home in both the PISA reading literacy and JCE English models. In the PISA model, effects are greater for male students with up to 100 books in the home, while effects are somewhat greater for female students with more than 100 books. In the JCE English model, the effect of books in the home is greater for female than for male students at all levels, except for zero books. The two models confirm the importance of books in the home to achievement, and underline the need to examine the relationship in greater detail as it may represent an important policy lever.

The final models for JCE Mathematics and Science are broadly similar to those for PISA mathematical and scientific literacy. However, whereas the model of PISA mathematical literacy includes an interaction between gender and lone-parent status (with a negative effect for female students in lone-parent families, and no difference in the effect of living/not living in a lone-parent family for males), the model for JCE Mathematics includes a gender interaction with frequency of completion of homework on time and no interaction between gender and lone-parent status. Both the PISA scientific literacy and JCE Science models include an interaction between gender and the index books in the home. In the PISA scientific literacy, the effects for male students with fewer than 250 books are greater than for females. In JCE Science, the effects for female students are greater than for male students for levels of more than 50 books. The variable number of siblings is included in the final models for both PISA scientific literacy and JCE Science.

A consideration of the outcomes across the final PISA and Junior Certificate Examination models leads to the following observations. First, while the estimated effects for student socioeconomic status as a stand-alone variable are significant in all six models, they are generally small. Other variables with significant effects, including home processes and attitudinal variables, are also important, though some of these, such as number of books in the home, are likely to be associated with socioeconomic status. Second, with the exception of the model for PISA reading literacy, where there was a positive effect for female students of about one-third of a standard deviation (32 points), gender differences could only be evaluated in relation to their interactions with other variable. An interaction between gender and the index of the number of books in the home was observed in three of the six final models. In the case of JCE English, for example, the contribution to the fitted scores of male students with 251-500 books is over one-half of a standard deviation (one grade), while the contribution to the scores of females is nine-tenths of a standard deviation (about one and a half grades). Third, the two school-level variables, school type and disadvantaged status, have significant effects in all six models. The effects associated with disadvantaged status, for example, range from one-quarter to one-third of a standard deviation. Such effects are non-trivial, given that student-level socioeconomic status is also found in each of the six models. This finding supports efforts designed to increase the educational achievement of students in schools designated as disadvantaged. The negative effects associated with vocational schools in all six models raise issues about the allocation of students to schools. Fourth, the two student-level categorical variables, frequency of absence from school and frequency of completion of homework on time, are found in all six final models, though in the JCE Mathematics model, the effect of completion of homework on time must be interpreted in terms of its interaction with gender. While the effects of absence from school and completion of homework are relatively small, they represent aspects of learning over which students can exercise some level of control (in contrast with variables such as socioeconomic status, which are less amenable to change). Fifth, the effect of student dropout risk is significant in all six final models. It functions as a random coefficient in the models for PISA reading literacy and JCE English, indicating that the strength of the association of perceived dropout risk with achievement varies across schools. Sixth, the effect of the variable attitude to reading is significant in the models for PISA reading literacy and JCE English (where it also interacts with gender). The role of attitude could be explored further to identify why it has a relatively strong association with achievement in reading/English, and consideration might be given to ways in which positive attitudes to reading could be fostered among students. Finally, the finding that higher levels of 'frequency of leisure reading' have negative effects in the PISA reading literacy and JCE English models is worth examining further. Future research might make a distinction between the volume of leisure reading in which students engage, and their overall engagement in reading at home and at school.

REFERENCES

Agresti, A., & Finlay, B. (1997). *Statistical methods for the social sciences* (3rd ed.). Boston: Prentice Hall.

Aitkin, M., Anderson, D., Francis, B. & Hinde, J. (in press). *Statistical modelling in GLIM* (2nd ed.). Oxford: Oxford University Press.

Kellaghan, T., & Dwan, B. (1995). *The 1994 Junior Certificate Examination: A review of results.* Dublin: National Council for Curriculum and Assessment.

Kellaghan, T., Madaus, G.F., & Rakow, E.A. (1979). Within-school variance in achievement: School effects or error? *Studies in Educational Evaluation, 5,* 101-107.

Martin, M.O., & Hickey, B.L. (1993). *The 1992 Junior Certificate Examination: A review of results.* Dublin: National Council for Curriculum and Assessment.

Martin, M.O., Mullis, I.V., Gregory, K.D., Hoyle, C., & Shen, C. (2000). *IEA's Third International Mathematics and Science Study: Effective schools in science and mathematics.* Boston: TIMSS International Study Center, Boston College.

OECD (Organisation for Economic Co-operation and Development) (2001). *Knowledge and skills for life: First results of PISA 2000.* Paris: Author.

Postlethwaite, T. N. (1995). Calculation and interpretation of between-school and within-school variance in achievement. In OECD, *Measuring what students learn* (pp. 81-91). Paris: OECD.

Raudenbush, S.W., Cheong, Y.F., & Fotiu, R.F. (1994). Synthesising cross-national classroom effects data: Alternative models and methods. In M. Binkley, K. Rust, & M. Winglee (Eds.), *Methodological issues in comparative educational studies: The case of the IEA Reading Literacy Study* (pp. 243-285). Washington, DC: US Department of Education, National Center for Educational Statistics.

Shiel, G., Cosgrove, J., Sofroniou, N., & Kelly, A. (2001). *Ready for life: The literacy achievements of Irish 15-year olds with comparative international data.* Dublin: Educational Research Centre.

Smyth, E. (1999). *Do schools differ? Academic and personal development among pupils in the second-level sector.* Dublin: ESRI/Oak Tree Press.

Snijders, A.B., & Bosker, R.J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modelling.* London: Sage.

APPENDIX 1

CANDIDATE VARIABLES FOR HIERARCHICAL LINEAR MODELS OF PERFORMANCE ON JUNIOR CERTIFICATE EXAMINATION ENGLISH, MATHEMATICS, AND SCIENCE

Variable (Level)	*Description*	*English*	*Mathematics*	*Science*
Level 2 (School)				
School Size/Stratum	Categorical (3): Large (81+ 15–year olds); Medium (41–80); Small: 17–40); Reference category: Medium	✓	✓	✓
School Sector	Categorical (3): Secondary, Community/ Comprehensive, Vocational; Reference category: Vocational	✓	✓	✓
Disadvantaged Status	Binary (Yes/No); Reference category: Yes	✓	✓	✓
Gender Composition	Categorical (3): All boys; All girls; Mixed. Reference category: Mixed	✓	✓	✓
Negative disciplinary Climate	Continuous; Composite/Weighted Likelihood Estimate (WLE) based on student responses to statements about behaviour in class; OECD Mean = 0.0; SD = 1.00	✓	✓	✓
Student–Teacher Ratio	Continuous (total enrolment divided by number of teachers); Mean = 15.1; SD = 1.82	✓	✓	✓
Level 1 (Student)				
Gender (Male)	Binary (Male, Female); Reference category: Female	✓	✓	✓
Socioeconomic status (SES)	Continuous - Range: 16 (Min) to 88 (Max);	✓	✓	✓
Parents' Education (Par Ed)	Categorical (4): Primary; Lower Secondary; Upper Secondary; Third Level; Reference category: Upper Secondary	✓	✓	✓
Lone Parent Status (Sing Par)	Binary (Yes, No); Reference category: Yes	✓	✓	✓
Number of Siblings (No. Siblings)	Continuous: 0–11; Mode: 2	✓	✓	✓
Parent Engagement (Par. Engage)	Continuous, WLE; OECD Mean = 0.0; SD = 1.00	✓	✓	✓
Log (Index of Number of Books in the Home)	Categorical (7 categories)	✓	✓	✓
Dropout Risk	Binary (High, Low); Based on student's intention to drop out before L. Cert examination; Reference category: No	✓	✓	✓
Absence from School	Categorical (3): 0 Absences (No absences in two weeks prior to PISA); 1–2 Absences; 3+ Absences; Reference category: 1–2 Absences	✓	✓	✓

APPENDIX 1 - Continued

Variable (Level)	*Description*	*English*	*Mathe-matics*	*Science*
Level 1 (Student)				
Completion of Homework on Time	Categorical (4): No homework (Homework never done on time); Sometimes (Sometimes done on time); Mostly (Mostly done on time); Always (Always done on time); Reference category: Mostly	✓	✓	✓
Junior Certificate Year	Categorical: 1999 or 2000 Reference category: 1999	✓	✓	✓
Diversity of Reading	Continuous: WLE based on frequency with which students read six types of texts; OECD Mean = 0.0; SD = 1.00	✓		
Frequency of Reading for Enjoyment	Categorical (4): none = no reading; 30 = fewer than 30 minutes per day; 30–60 = 30–60 minutes per day; 60+ = more than 60 minutes; Reference category: Up to 30 minutes per day	✓		
Attitude Towards Reading	Continuous; WLE based on student responses to 9 statements about attitude towards reading; attitude; OECD Mean = 0.0; SD = 1.00	✓		

Note. A tick indicates that the variable was tested for inclusion in the models for English, Mathematics and/or Science.