The Irish Journal of Education, 2006, *xxxvii,* pp. 5-26.

AN OVERVIEW OF THE MAIN FINDINGS OF PISA 2003 IN IRELAND

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

In PISA 2003, 15-year olds in Ireland achieved a mean score on the major assessment domain, mathematics, that is not significantly different from the OECD country average. Irish mean scores on two minor domains, reading literacy and science, are significantly higher than the OECD country average scores, while performance on a third, cross-curricular problem solving, is not significantly different. Several reasons for the relatively poor performance of students in Ireland on the mathematics scale are considered, including differential performance on the mathematics subdomains, poor performance among higher-achieving students, and differences between the Junior Certificate mathematics examination and PISA mathematics in terms of the contexts in which PISA items were embedded, and the content they tapped. School and student variables associated with achievement on PISA mathematics are presented descriptively, and in the context of a multi-level model of achievement. The model, which explained 79% of between-school variance, and 30% of within-school variance, summarizes the contributions of school socioeconomic status, school disciplinary climate in mathematics lessons, student attendance at school, number of books in the home, home educational resources, and gender.

PISA 2003 is the second in a series of surveys of the educational achievements of 15-year olds organized by member-countries of the OECD. In 2003, mathematics was the major assessment domain, while reading literacy, science, and problem solving were minor domains. Since PISA adopts a 'literacy-based' approach to assessment, students were asked to apply skills and knowledge to authentic, 'real-life' problems that were deemed to be important for their future lives, rather than engage with prescribed curriculum content. Hence, unlike examinations such as the Junior Certificate, PISA does not provide a direct assessment of school-based learning. Indeed, according to OECD (1999):

> Although the domains of reading literacy, mathematical literacy and scientific literacy correspond to school subjects, the OECD assessments will not primarily examine how well students have mastered the specific curriculum content. Rather, they aim at assessing the extent to which young people have acquired the wider knowledge and skills in these domains that they will need in adult life. (p.9)

A 'literacy' approach is justified on the basis of two main principles. First, that of breadth: it allows for a wider range of knowledge and skills to be assessed. Second, relevance: it is forward-looking, anticipating the knowledge and skills that students are likely to need and build on in adult life, rather than backward-looking, which would entail an assessment of what has already been learned in school.

The PISA tests are based on underlying assessment frameworks that have been established by international teams of experts and accepted by participating OECD countries (see OECD, 2003).

PISA implements a cross-sectional design. In 2003, representative samples of 15-year olds in 30 OECD member countries and in 11 partner countries were asked to complete paper-and-pen tests and a Student Questionnaire at a given point in time. PISA also allows for a comparison of achievement over time, as there is built-in overlap between tests administered in different cycles. Hence, in 2003, it was possible to compare performance with performance in 2000.

PISA uses a rotated booklet design where each student responds to a single booklet. In 2003, there was a total of 13 booklets. All booklets included some mathematics items, while 7 included reading literacy items, 7 included science items, and 7 included problem solving items. Booklets were assigned at random to students. Item response models were used to develop achievement scales, with item difficulties placed on the same scale as student scores. In Ireland, 139 schools and 3,880 students took part. The weighted school response rate in Ireland was 92.8% after replacements and the corresponding student response rate was 82.6 percent. Both school- and student-level response rates exceeded OECD requirements.

In scaling the PISA 2003 data, the scores of students who were not asked to attempt items in reading, science, and problem solving were imputed, using their scores in mathematics and other relevant background characteristics. Hence, unlike PISA 2000, scores in all four domains were available for each student who participated in PISA 2003.

MAIN ACHIEVEMENT OUTCOMES

This section describes performance in mathematics, reading literacy, science, and cross-curricular problem solving in PISA 2003. Following this, between-school variance in achievement is described.

Mathematics – Combined Scale

On the combined mathematics scale, students in Ireland achieved a mean score of 502.8 (SE = 2.45). This score is not significantly different from the OECD country average (500.0; SE = 0.43), and ranked Ireland 17th among 29 OECD countries (95% confidence interval for Ireland's ranking = 15th to 18th), and 20th among participating countries (95% confidence interval: 17th to 21st). Students in Hong-Kong China, the highest-scoring country, achieved a mean score (550.4) that was one-half of a standard deviation higher than the mean score of students in Ireland. Countries with overall mean scores that did not differ significantly from Ireland's were the Czech Republic, Denmark, France, Sweden, Austria, Germany, the Slovak Republic, and Norway. In Ireland, the standard deviation was 85.3 points. Only Finland among OECD countries has a lower standard deviation (Table 1)[1] (Cosgrove, Shiel, Sofroniou, Zastrutzki, & Shortt, 2005).

Performance on combined mathematics was also reported with reference to proficiency levels. These are descriptive summaries of the types of tasks that students at different points on the scale are likely to succeed on. Table 2 gives the percentages of students in Ireland and across OECD countries that scored at each proficiency level. The percentage of students in Ireland scoring below Level 1 (4.7%) is lower than the corresponding OECD country average (8.2%). On the other hand, fewer students in Ireland (11.3%) perform at Levels 5 and 6 compared with the OECD average (16.5%). In countries with overall performance that was significantly higher than Ireland's, more students tend to score at Levels 5 and 6 (e.g., 23.4% in Finland, 24.8% in Korea, and 25.5% in the Netherlands), and fewer score below Level 1 (e.g., 1.5% in Finland, 2.5% in Korea, and 2.6% in the Netherlands).

Mathematics – Subscales

Performance on PISA 2003 mathematics can be described with reference to four mathematics subscales: Change & Relationships, Quantity, Space & Shape, and Uncertainty. Students in Ireland achieved a mean score that is significantly above the OECD country average on two scales – Change & Relationships (506.0 vs. 498.8), and Uncertainty (517.2 vs. 502.0); that is not significantly different on one scale – Quantity (501.7 vs. 500.7); and is significantly lower on one – Space & Shape (476.2 vs. 496.3). Among 29 OECD countries, Ireland's rankings ranged from 10th on Uncertainty to 23rd on Space & Shape.

[1] The achievement data for the United Kingdom were not included in multiple-comparison tables because its school- and school-level responses rates were below the required standards. Analysis suggested potential non-response bias at student level.

Table 1

Mean Achievement Scores and Standard Deviations on Combined Mathematics – OECD and Partner Countries

	Mean	(SE)	SD	OECD Diff		Mean	(SE)	SD	OECD Diff
Hong Kong-Ch	550.4	(4.54)	100.2	▲	Norway	495.2	(2.38)	92.0	▼
Finland	544.3	(1.87)	83.7	▲	Luxembourg	493.2	(0.97)	91.9	▼
Korea	542.2	(3.24)	92.4	▲	Poland	490.2	(2.50)	90.2	▼
Netherlands	537.8	(3.13)	92.5	▲	Hungary	490.0	(2.84)	93.5	▼
Liechtenstein	535.8	(4.12)	99.1	▲	Spain	485.1	(2.41)	88.5	▼
Japan	534.1	(4.02)	100.5	▲	*Latvia*	483.4	(3.69)	87.9	▼
Canada	532.5	(1.82)	87.1	▲	United States	482.9	(2.95)	95.2	▼
Belgium	529.3	(2.29)	109.9	▲	*Russian Fed.*	468.4	(4.20)	92.3	▼
Macao-Ch	527.3	(2.89)	86.9	▲	Portugal	466.0	(3.40)	87.6	▼
Switzerland	526.6	(3.38)	98.4	▲	Italy	465.7	(3.08)	95.7	▼
Austria	524.3	(2.15)	95.4	▲	Greece	444.9	(3.90)	93.8	▼
New Zealand	523.5	(2.26)	98.3	▲	*Serbia & Monte.*	436.9	(3.75)	84.7	▼
Czech Rep	516.5	(3.55)	95.9	▲	Turkey	423.4	(6.74)	104.7	▼
Iceland	515.1	(1.42)	90.4	▲	*Uruguay*	422.2	(3.29)	99.7	▼
Denmark	514.3	(2.74)	91.3	▲	*Thailand*	417.0	(3.00)	82.0	▼
France	510.8	(2.50)	91.7	▲	Mexico	385.2	(3.64)	85.4	▼
Sweden	509.0	(2.56)	94.7	▲	*Indonesia*	360.2	(3.91)	80.5	▼
Austria	505.6	(3.27)	93.1	o	*Tunisia*	358.7	(2.54)	82.0	▼
Germany	503.0	(3.32)	102.6	o	*Brazil*	356.0	(4.83)	99.7	▼
Ireland	**502.8**	**(2.45)**	**85.3**	o	OECD Total	489.0	(1.07)	103.6	
Slovak Rep	498.2	(3.35)	93.3	o	OECD Average	500.0	(0.63)	100.0	

Legend (shading): dark = Mean significantly higher than Ireland; white = Mean not significantly different from Ireland; light grey = Mean significantly lower than Ireland

▲ Above OECD average
o At OECD average
▼ Below OECD average

The column "OECD Diff" indicates whether each country scores at, significantly above, or significantly below the OECD average (p<.05), using Bonferroni-adjustments with an overall alpha-level of .05.

Source: Cosgrove et al. (2005), Table 3.1.

Table 2
Percentages of Students Achieving at Each of Six Mathematics Proficiency Levels – Ireland and OECD Country Average

Level	Summary Description	Ireland % (SE)	OECD % (SE)
Level 6	Evaluating, generalizing and utilizing information from mathematical investigations; modelling of complex problem situations.	2.2 *(0.33)*	4.0 *(0.1)*
Level 5	Developing and working with mathematical models of complex situations.	9.1 *(0.76)*	10.6 *(0.13)*
Level 4	Working with mathematical models of concrete situations.	20.2 *(1.06)*	19.1 *(0.17)*
Level 3	Solving problems in familiar contexts that require multiple steps for solution.	28.0 *(0.82)*	23.7 *(0.18)*
Level 2	Solving problems that require no more than direct inference.	23.6 *(0.83)*	21.1 *(0.15)*
Level 1	Performs clearly-defined tasks with familiar contexts where all relevant information is present and no inference is required.	12.1 *(0.84)*	13.2 *(0.16)*
Level 1	Has less than .50 chance of responding correctly to Level 1 tasks. Mathematics achievement not assessed by PISA.	4.7 *(0.57)*	8.2 *(0.17)*

More detailed descriptions of the mathematics proficiency levels can be found in OECD (2004) and Cosgrove et al. (2005).
Source: Cosgrove et al. (2005), Table 3.11.

Comparison between Performance on PISA 2000 and PISA 2003 Mathematics

Performance on two subscales – Change & Relationships and Space & Shape – can be compared between 2000 and 2003. Although the OECD country average score on Change & Relationships increased between 2000 and 2003, and mean performance improved in eight countries, performance in Ireland remained the same. Neither the OECD country average, nor the mean score of students in Ireland, changed on the Space & Shape scale between 2000 and 2003.

Reading Literacy

In PISA 2003, as reading literacy was a minor assessment domain, it was assessed using a smaller pool of items than in PISA 2000, and performance was reported for a combined scale only. Ireland's mean score in 2003 was 515.5, which is significantly higher than the OECD country mean of 494.2. Ireland ranked 7th of 40 participating countries (95% CI = 6th to 10th), and 6th of 29 OECD countries (95% CI = 6th to 8th). Three countries, Finland (Mean = 543.5), Korea (534.1) and Canada (527.9), had mean scores that are significantly higher than Ireland's. Eleven percent of students in Ireland achieved at the lowest levels (Level 1 and below) on the combined reading proficiency scale, compared with an OECD average of 19.1%, while 35.5% achieved at the highest levels (Levels 5 and 6), compared with an OECD average of 29.6 percent. In the highest scoring country, Finland, just 5.7% achieved at Level 1 or below, while 48.1% achieved at Levels 5 and 6. Ireland was one of three OECD countries that registered a significant decline in performance on combined reading literacy in 2003 relative to 2000. Ireland's score fell from 526.7 to 515.5, with students scoring at the 75th, 90th, and 95th percentiles doing less well in 2003 than in 2000.

Science

As PISA science was a minor domain in both 2000 and 2003, performance was reported only on an overall scale in both years. In 2003, students in Ireland achieved a mean score of 505.4, which is significantly higher than the OECD country average (499.6). Ireland ranked 16th of 40 countries (95% CI = 12th to 20th) and 13th of 29 OECD countries (95% CI = 9th to 16th). The scores of students in Ireland at the 5th, 10th, and 25th percentiles are higher than the corresponding OECD country average scores; the score at the 75th percentile is not significantly different; and the scores at the 90th and 95th percentiles are significantly lower.

Neither the OECD country average score, nor the mean score for students in Ireland, changed in science between 2000 and 2003. However, the OECD average scores at the 5th and 10th percentiles declined between the two years, and the OECD average scores at the 90th and 95th percentiles increased. In Ireland, no differences were observed at any of these benchmarks.

Cross-Curricular Problem Solving

In 2003, a test of cross-curricular problem solving was included in PISA. Items on the test assessed reasoning skills, including analytical and analogical reasoning. The problems were intended to be independent of other assessment domains, such as mathematics. Students in Ireland achieved a mean score of 498.5, which is not significantly different from the OECD average of 500.0. Ireland ranked 21st of 40 participating countries (95% CI = 20th to 22nd), and 18th of 29 OECD countries (95% CI = 17th to 19th). Students in Ireland scoring at the 5th, 10th, and 25th percentiles achieved scores that are significantly higher than the corresponding OECD country average scores. However, students in Ireland scoring at the 75th, 90th, and 95th percentiles achieved scores that are significantly lower. The correlation between performance on PISA 2003 mathematics and problem solving is .90 (t = 74.61, p .001). This is marginally stronger than the correlation between reading literacy and problem solving (.87, t = 82.70, p= .001), and between science and problem solving (.85, t = 143.77, p= .001).

Between-School Variance in Achievement

In Ireland, the intra-cluster correlation (ICC), or the percentage of variance that is between schools, was 16.7 for combined mathematics, 22.5 for reading literacy, 16.2 for science, and 15.7 for problem solving. In mathematics, Ireland's ICC is eighth lowest among 29 OECD countries (mean:32.7%), and among 39 countries in all. While some of the Scandinavian countries had between-school variance that was less than 10%, in eight countries, including Belgium, Italy, Germany, and Austria, between-school variance exceeded 50 percent. In this context, the ICC for Ireland can be interpreted as indicating relatively small differences between schools in achievement. In the three minor domains, between-school variance for Ireland is also below the corresponding OECD country average (31.4% for reading literacy; 29.9% for science; and 31.6% for problem solving).

VARIABLES ASSOCIATIED WITH ACHIEVEMENT ON PISA 2003

Questionnaires completed by students and their principal teachers provided data for a wide range of student- and school-level variables, which were used to guide the interpretation of performance outcomes. In this section, descriptive statistics, including correlation coefficients, are used to describe associations between individual student and school variables and mathematics achievement. Following this, a hierarchical multi-level model is employed to estimate the contributions of selected variables to achievement, while controlling for the contribution of others.

Student-Level Variables

A significant gender differences in favour of male students on the combined mathematics scale was observed in 21 of 29 OECD countries. In Ireland, male students achieved a mean score that is 14.8 points (one-sixth of a standard deviation) higher than the mean score of female students. This is slightly larger than the OECD country average difference of 11.1 points. Irish male students outperformed their female counterparts on all four mathematics subscales, with the largest difference (25.5 points, or one-quarter of a standard deviation) observed for Space & Shape. In reading literacy, female students outperformed males by 29 score points (one-third of a standard deviation). Small differences in favour of male students in Science (2.0 points) and Problem Solving (0.5 points) are not statistically significant.

Student socioeconomic status (based on the highest occupation of either parent) was also associated with achievement. Students in the top third of the distribution of SES scores achieved a mean score on combined mathematics that is 30 points (almost one-third of a standard deviation) higher than the mean score of students in the middle third, and 62 points (two-thirds of a standard deviation) higher than the mean score of students in the bottom third. Similar patterns were observed for the other content domains.

In all four domains, students' current grade level was associated with achievement. On combined mathematics, for example, students in third year outperformed students in second year by 85 points (almost one standard deviation), but did less well than students in transition year (by 50.6 points) and students in fifth year (by 22.8 points).

Other student-level variables associated with achievement in one or more domains include number of siblings, home educational resources,[2] number of books in the home, frequency of absence from school, mathematics self-efficacy, anxiety about mathematics, frequency of reading fiction, and frequency of reading e-mails/web pages (see Table 3). A notable feature of Table 3 is the stability of correlation coefficients across the PISA assessment domains, with the possible exception of the correlations involving home educational resources. For example, the association between socioeconomic status and achievement ranges from .32 (mathematics) to .34 (science and problem solving). The strongest correlation is between mathematics self-efficacy and achievement. However, since the items on the PISA 2003 student questionnaire that were designed to measure self-efficacy asked students about their confidence in solving specific mathematical tasks (e.g., calculating how much cheaper a television set would be after a 30% discount, or solving an equation for x where $3x + 5 = 17$), this variable could be interpreted as an outcome of learning. The negative correlation between anxiety about mathematics and achievement in mathematics indicates that, as anxiety increases, performance tends to decline.

[2] Home educational resources was a composite variable based on whether a student had one or more of the following: a desk for study, a quiet place to study, books to help with school work.

Table 3
Correlations between Selected Student Variables and Achievement

Variable	Correlation with Mathematics	Correlation with Reading Literacy	Correlation with Science	Correlation with Problem Solving
Socioeconomic status	.315*	.329*	.337*	.336*
Number of siblings	-109*	-.136*	-.123*	-.120*
Home educational resources	.231*	.406*	.263*	.223*
Number of books in the home	.335*	.353*	.365*	.348*
Absence from school	-.185*	-.190*	-.186*	-.175*
Mathematics self-efficacy	.529*	-	-	.469*
Mathematics anxiety	-.363*	-	-	-.325*
Frequency of reading fiction	-	.285*	-	-
Freq of reading e-mails/ web	-	.207*	-	-

* p = .001; df = 80 [number of variance strata associated with the balanced repeated replicate (BRR) method of variance estimation].
Source: Cosgrove et al. (2005), Table 4.36.

School-Level Variables

A number of school-level variables were also shown to be related to student achievement. Variables were based on data in the School Questionnaire (e.g., computer-student ratio), the Student Questionnaire (e.g., disciplinary climate in mathematics classes)[3], and the Department of Education and Science database of post-primary schools (e.g., school sector, school disadvantaged status).

Students attending secondary schools achieved significantly higher mean scores in all four assessment domains than students attending other schools. For example, on combined mathematics, students in secondary schools achieved a mean score that was 16.9 points (one-fifth of a standard deviation) higher than students attending community/comprehensive schools, and 40.1 points (almost one-half of a standard deviation) higher than students attending vocational schools.

Students in all-boys' schools outperformed students in all-girls' schools by 22.4 points (over one-quarter of a standard deviation) in mathematics, and by 11.8 points (one-eighth of a standard deviation) in science. Students in all-girls'

[3] Disciplinary climate in mathematics classes was first aggregated to the school level, and then disaggregated to the student level, so that each pupil was assigned the average value for his/her school.

schools outperformed students in all-boys' schools by 17.7 points (one-fifth of a standard deviation) in reading literacy.

Associations between other school-level variables and achievement are presented in Table 4. The variable Junior Certificate examination (JCE) fee waiver represents the proportion of students in a school who were eligible for a waiver (usually on the basis that they or their parents had a medical card). This variable can be interpreted as a measure of school-level socioeconomic status. Correlations between it and achievement range from -.29 (reading literacy) to -.32 (science). The associations between disciplinary climate in mathematics classes and achievement are weaker, ranging from .19 (science) to .23 (reading literacy).

Table 4
Correlations between Selected School Variables and Achievement

Variable	Correlation with Mathematics	Correlation with Reading Literacy	Correlation with Science	Correlation with Problem Solving
Junior Cert. fee waiver	-.308*	-.287*	-.317*	.306*
Percent of 15-year olds female	-.064*	.103*	-.015	
Ratio of computers to students	-.185*	-.221*	-.194*	
Total minutes of instructional time per week	.170*	.165*	.145*	
Disciplinary climate in mathematics classes	.211*	.227*	.188*	.190*

* p = .001; df = 80 [number of variance strata associated with the balanced repeated replicate (BRR) method of variance estimation].
Source: Cosgrove et al. (2005), Table 4.38.

A Multi-Level Model of Achievement in PISA Mathematics

Hierarchical linear models for student achievement in mathematics, reading literacy, and science were developed (see Goldstein, 1995; Longford, 1993; Snijders & Bosker, 1999). These are linear regression models with random components at the cluster (school) and individual levels. The variation in the intercept that occurs from school to school is fitted with a random effect that follows a Normal distribution. The residual (within-school) variation is also fitted with a Normal distribution. The estimated effects of variables at the student level are allowed to vary across clusters (schools) by including random effects for their parameters, known as random coefficients. This suggests a range of values that the parameter estimate takes over the population of clusters (schools), with the fitted Normal distribution for the random effect. The NLME library of Pinheiro and Bates (2000), implemented in the R statistical package, was used for fitting the multi-level models. Models were estimated by full maximum likelihood, which allows deviance tests of the significance of both fixed and random effects to ascertain the fit of the model, with degrees of freedom set to the difference in the number of fitted terms (McCullagh & Nelder, 1989). The multi-level model for combined mathematics is described below.

A broad range of student- and school-variables were considered for modelling purposes. First, each of several student- and school-level variables was fitted separately. All of these separate models provided statistically significant improvements over the null model. Eight student-level variables (gender, socioeconomic status, lone-parent status, number of siblings, number of books in the home[4], home educational resources, frequency of absence from school, and grade level) and four school-level variables (size, sector, disciplinary climate in mathematics classes, and percent of students in the school who received a Junior Certificate examination fee waiver) were candidates for inclusion in the final model. All eight student-level variables remained in the final model, along with two of the four school-level variables (disciplinary climate, percent JCE fee waiver). In addition to main effects, there was one student-level interaction (log of books index by frequency of absence from school) (Table 5). A sub-level model that fits student-level variables and interactions alone explained 62.3% of the variance at cluster (school) level[5], and 26.9% at individual (class and student) level. Including the school-level variables improved explanation by an additional 16.5% and 2.7% respectively. The final model explained 78.8% of between-cluster (school) variance, and 29.6% of within-cluster (classes and students) variance.

[4] Books in the home consisted of the following index: 0-10; 11-25; 26-100; 101-200, 201-500; 500+.

[5] The ICC of the null random-intercept model is 14.6%. This differs from the estimate of 16.7% referred to above, as it is based on students for whom complete data for all variables in the model were available, and is also linked to the use of full maximum likelihood estimation in modelling.

Table 5
Final Multi-Level Model of Achievement in Mathematics

	Parameter	*SE*	*Test Statistic*	*df*	*p-value*
Intercept	473	6.575			
Student-Level Main Effects					
Gender: female-male	-24.233	2.745	t=-8.824	42082	<.001
Socioeconomic status	0.794	0.090	t=8.868	263	<.001
Lone parent: dual-lone	15.670	3.448	t=4.544	5498	<.001
Number of siblings			Ddiff=25.328	4	<.001
none-one	-14.708	4.707			
two-one	-9.242	3.620			
three-one	-13.218	4.041			
four or more-one	-17.842	4.090			
Log books index	38.745	3.837			
Home educational resources			Ddiff=15.207	2	<.001
low-high	-11.154	3.915			
medium-high	-8.763	2.905			
Absence					
1 or 2 days-no days	3.533	7.025			
3 days or more-no days	-16.534	10.658			
Grade level			Ddiff=276.094	3	<.001
Grade 8-Grade 9	-59.191	7.874			
Grade 10-Grade 9	33.917	4.118			
Grade 11-Grade 9	35.654	3.560			
School-Level Main Effects					
Disciplinary climate	19.224	4.451	t=4.319	142	<.001
Percent JCE fee waiver	-0.8	0.121	t=-6.594	142	<.001
Student-Level Interactions					
Log (books index) x Absence			Ddiff=7.355	2	.025
Log (books index) x 1 or 2	-12.615	5.737			
Log (books index) x 3 or more	-15.069	8.534			
Variance Components					
Intercept	172.169				
Residual	4722.511				

Source: Cosgrove et al. (2005), Table 5.4.

The value of the intercept (473.0 score points) represents the score for a hypothetical student who is male, of average SES, in a lone-parent family, with one sibling, with zero to ten books in the home, with high home educational resources, with no absence in the fortnight prior to the PISA assessment, in grade 9 (third year), attending a school with an average level of disciplinary climate in mathematics classes, and with an average level of students receiving a fee waiver for the Junior Certificate examination. The insertion of the contribution of other fitted scores (calculated using the parameter estimates) produces fitted scores for different values of the student and school variables. Unlike continuous variables (e.g., socioeconomic status)[6], the values of categorical variables can be read directly from Table 5. Thus, adjusting for other variables in the model, the gender difference represents a deficit of 24.2 points or about three-tenths of a standard deviation for females. This can be contrasted with the difference of 14.4 points referred to above, when the association between gender and mathematics achievement was considered without reference to any other variables. In the final model, a student from a dual-parent family scores 15.7 points (one-sixth of a standard deviation) higher than a student from a lone-parent family. A student who is in grade 8 scores on average 59.2 points lower than a student in grade 9, while students in grades 10 and 11 score higher by values of 33.9 and 35.7 respectively.

Table 6 gives the raw scores and fitted values for three continuous variables, student socioeconomic status, JCE fee waiver (a measure of school-level socioeconomic status), and school disciplinary climate for mathematics. In the case of student-level socioeconomic status, compared to a student at the mean of the high group of the SES variable (14.4 score points), students at the mean of the medium group are expected to score 14.9 points (about one-sixth of a standard deviation) lower,[7] and students at the mean of the low group 28.6 points (about one-third of a standard deviation) lower. Compared to students in schools at the mean of the high JCE fee waiver rates (–14.4), students in schools at the mean of the medium and low levels score 16.6 and 26.5 points higher (one-sixth of a standard deviation and one-third of a standard deviation, respectively). Relative

[6] The values for continuous variables which appear linearly as a main effect only can be interpreted as the change in achievement score corresponding to a one-unit increase in the explanatory variable. To calculate actual values, the parameter estimates for continuous variables are first multiplied by the chosen example value of the variate (first subtracting the mean of the variable if it was centred during model fitting).

[7] This is obtained by subtracting the fitted value for the medium group from the fitted value for the high group.

to students in schools at the mean of the high level of school disciplinary climate (26.1 points), students in schools at the medium and low levels score 25.5 and 52.1 points higher (three-tenths and three-fifths of a standard deviation) respectively.

Table 6
Contribution to Fitted Scores in Combined Mathematics for Example Values of Selected Continuous Variables

	Low	Medium	High
Student Socioeconomic Status			
Raw Score	30.54	47.88	66.59
Contribution to Student Score	–14.26	–0.48	14.37
School JCE Fee Waiver			
Raw Score	10.67	23.07	43.80
Contribution to Student Score	12.14	2.21	–14.37
School Disciplinary Climate			
Raw Score	–1.09	0.29	1.62
Contribution to Student Score	–26.03	0.56	26.07

Source: Cosgrove et al. (1005), Tables 5.5, 5.6, and 5.7.

The contributions of fitted scores for the interaction of frequency of absence from school and index of books in the home are listed in Table 7. The impact of books in the home levels off as the upper value is approached, with the greatest relative increase occurring for the no absence condition. There is little difference between the none and 1-2 days absence score contributions for students with less than 11 books in the home (0.0 vs. 3.5 points), but there is a large deficit for three or more days absence (-16.5 points). The large effect of the book index is confirmed for all students. For example, there is a difference of 69.4 points (four-fifths of a standard deviation) between zero to 10 and 500+ books for students with no absence, and a corresponding difference of 42.4 points (half a standard deviation) for students absent 3 or more days.

Table 7
Contribution to Fitted Scores in Combined Mathematics for Books in the Home by Frequency of Absence from School

	Books Index					
Absence	None to 10	11-25	26-100	101-200	201-500	500+
None	0.0	26.86	42.57	53.71	62.36	69.42
1 or 2 days	3.53	21.64	32.24	39.76	45.59	50.35
3 or more days	–16.53	–0.12	9.48	16.29	21.57	25.89

Source: Cosgrove et al. (2005), Table 5.8.

Finally, two variables, anxiety towards mathematics and self-efficacy in mathematics, were added to a model of combined mathematics (see Cosgrove et al., 2005, Appendix B, Sec. B1). A difficulty with these variables is that they affect and are affected by students' recent academic performance. Hence, although there was an improvement in explained variation of 4.9% at the cluster (school) level, and 14.7% at the student level over that of the final model, giving resultant values of 83.8% and 44.3% of explained variance respectively, the interpretation of a model that includes these variables is complex, given the possible status of anxiety and self-efficacy as joint outcomes.

LINKING PISA TO NATIONAL CURRICULA

A value judgment underpins the PISA approach, insofar as international experts have identified the skills that they feel are relevant for successful participation in future learning and adult life. However, while this approach may give breadth and relevance to the assessment, it can make results difficult to interpret with respect to national curricula. One cannot assume uniform relevance across all countries and test items.

If national curricula differ substantially from what PISA assesses, a logical tension arises: how are the results supposed to establish benchmarks for educational improvement and help countries to understand (curricular or instructional) strengths and weaknesses of the system? As put by one author:

> There is a curious contradiction in the design of PISA. It is intended to be a knowledge base for policy analysis. Yet, it explicitly rejects attempting to assess what pupils have learned in relation to the school curriculum. This puts the onus on PISA to demonstrate that non-curriculum based tests can

> be used to derive policy conclusions for educational systems. (Smithers, 2004, p. 38)

Smithers argues that cross-country differences in degree of curriculum match can be a source of bias in PISA, and that lack of a curriculum match analysis severely limits PISA's explanatory power. Others have also commented on the tension between the PISA approach to assessment and claims that it can be used to draw inferences about the performance of education systems (Goldstein, 2004; Kaiser, Leung, Romberg, & Yashenko, 2002; Prais, 2003). This 'curious contradiction' points to the need to consider the relationship between PISA and national curricula.

PISA and the Curriculum in Ireland

PISA 2003 mathematics items were rated by three teachers with extensive experience in their subject area in terms of the expected familiarity of third year (grade 9) students with three aspects of the item: the concept, its context of application, and the item format. Between 50% and 70% of mathematics items were rated as somewhat or very familiar in terms of the concepts assessed (Table 8). Items that were rated as not familiar featured concepts relating to probability (not studied until senior cycle) and spatial visualization skills (not studied at all). The percentage of familiar concepts is higher for Higher level than for Ordinary level, and lowest for Foundation level. This is because Higher-level students study more topics in the areas of geometry, trigonometry, and algebra than Ordinary-level students, while the Foundation-level course tends to focus more on applied mathematics and a narrower range of topics. The majority of items (between 65% and 80%, depending on syllabus level) were rated not familiar in terms of the context in which the concept was embedded. This is because the vast majority of Junior Certificate examination mathematics questions are presented in pure mathematical contexts with little or no redundant information, whereas PISA embeds the mathematical information in real-life problem situations and the mathematics commonly has to be extracted from the context. The item format of most items was also rated as unfamiliar (between 62% to 85%), mainly because of the lack of the multiple-choice format on Junior Certificate mathematics examination papers.

Table 8
Number (and %) of PISA 2003 Mathematics Items Rated as Familiar in Concept, Context, and Format, by Junior Certificate Syllabus Level

	Not familiar		Somewhat familiar		Very familiar		Total	
	N	%	N	%	N	%	N	%
Concept								
Higher	26	30.6	21	24.7	38	44.7	85	100.0
Ordinary	30	35.3	25	29.4	30	35.3	85	100.0
Foundation	44	51.8	22	25.9	19	22.4	85	100.0
Context								
Higher	56	65.9	19	22.4	10	11.8	85	100.0
Ordinary	60	70.6	17	20.0	8	9.4	85	100.0
Foundation	68	80.0	14	16.5	3	3.5	85	100.0
Format								
Higher	53	62.4	21	24.7	11	12.9	85	100.0
Ordinary	62	72.9	17	20.0	6	7.1	85	100.0
Foundation	71	83.5	12	14.1	2	2.4	85	100.0

Source: Cosgrove et al. (2005), Table 6.13.

For PISA 2003, the Junior Certificate mathematics topic area in which the concept underlying each item featured was also identified. Most PISA items assessed concepts relating to the Junior Certificate mathematics topic of applied arithmetic and measure, and a number of items fell into the topic areas of statistics and number systems. However, no PISA mathematics items fell into the topic areas of sets, geometry, or trigonometry, which suggests a substantial divergence in national and international conceptualizations of the assessment of knowledge and skills in these mathematical areas.

Comparison of Performance on PISA and on the Junior Certificate Examination

In a comparison of performance on PISA 2003 and the Junior Certificate examination, almost three-quarters (72%) of Foundation-level students and over one-fifth (22%) of Ordinary-level students were found to have scored at or below Level 1 on the combined mathematics proficiency scale (Table 9), indicating that these students could, at best, complete familiar tasks where all relevant information is present, and carry out routine procedures (Cosgrove et al., 2005). The students are not considered to have the necessary mathematics skills to meet their future needs in real life, or to progress their education (OECD, 2004). On the other hand, 61% of Higher-level students achieved Level 4 or higher on the proficiency scale, indicating that a majority of students at this syllabus level could at least work effectively with explicit models for complex concrete situations linked directly to aspects of real-life situations.

Table 9

Percentage of Irish Students at Each Combined Mathematics Proficiency Level Cross-Tabulated with Junior Certificate Mathematics Syllabus Level

	Below Level 1		Level 1		Level 2	
	%	SE	%	SE	%	SE
Higher	0.3	0.16	1.2	0.33	9.0	1.09
Ordinary	4.1	0.70	17.8	1.26	36.2	1.27
Foundation	33.4	4.05	38.5	4.18	22.5	3.35

	Level 3		Level 4		Level 5	
	%	SE	%	SE	%	SE
Higher	28.8	1.26	35.8	1.52	19.7	1.39
Ordinary	30.4	1.30	9.9	0.90	1.5	0.38
Foundation	5.5	1.83	0.0	0.00	0.0	0.00

	Level 6	
	%	SE
Higher	5.2	0.76
Ordinary	0.1	0.11
Foundation	0.0	0.0

Total number of Irish students = 3880. N Higher = 1651; N Ordinary = 1941; N Foundation = 265; N Missing syllabus level = 24.

Source: Cosgrove et al. (2005), Table 6.19.

CONCLUSION

The overall performance in PISA 2003 mathematics of Irish students, earning them a ranking of 17th of 29 OECD countries, and 20th of 40 participating countries, and a mean score that is not significantly different from the OECD country average, indicates a level of performance that is in the average range by international standards. This performance may be contrasted with performance in reading and science, for which mean scores were above OECD country averages. There are a number of possible reasons for this. One relates to the relatively poor match between the content/processes of PISA and the Junior

Certificate syllabus/examination in mathematics. The contexts in which the majority of PISA items were embedded and the formats in which they were presented were deemed to be unfamiliar to Irish students. Another relates to the nature of teaching in schools which has been described as didactic, characterized by much drill and practice of mathematical procedures in a controlled setting, with little emphasis on the explanation of concepts, and providing few opportunities for students to engage in relational thinking, or in the application of mathematical knowledge in applied problem contexts (Lyons, Lynch, Close, Sheerin, & Boland, 2003). These findings suggest that greater attention may need to be paid to teaching methodologies and conceptual understanding at junior cycle level. There may be particular value in trying out and evaluating aspects of the PISA approach with students in Transition year, where there may be more time for exploration and discussion than at other levels.

Although there are some positive features associated with the performance of students in Ireland in mathematics – the comparatively low standard deviation and the small proportion of variance in achievement between schools – the performance of the highest achieving students must be a matter of concern. Students at the 75th, 90th, and 95th percentile ranks achieved scores that are significantly lower than the corresponding OECD country averages, while just 11% of students achieved scores at Levels 5 and 6 on the combined proficiency scale, compared with an OECD average of 15%, and 23% in Finland.

Similarities between the performance of students in Ireland on combined mathematics and cross-curricular problem solving is of interest, not least because the assessment of problem-solving was designed to be independent of mathematics. The strong correlation between the two domains ($r = .90$) suggests that a process such as analytical reasoning (which features strongly in the assessment framework for problem solving) is common to both mathematics and cross-curricular problem solving.

It is unclear if the apparent decline in mean performance in reading between 2000 and 2003 in Ireland, and in a number of other countries including Italy and Spain, arises for methodological reasons or is in fact real. In any event, the OECD cautions against inferring trends on the basis of performance at just two points in time (OECD, 2004).

The multi-level model of performance in mathematics in PISA 2003 confirms the contribution of school-level socioeconomic status to student achievement, while adjusting for other variables, including student-level socioeconomic status. The survival of school-level disciplinary climate in mathematics classes in the final model, after adjusting for school-level socioeconomic status, is noteworthy, and suggests that this variable, which is based on students' responses to questions about noise level in class, attentiveness of fellow students, and focus during lessons, is worth exploring in greater detail.

The model of mathematics achievement confirms the stronger performance of male students relative to female students, with a fitted score for males that is over one-quarter of a standard deviation higher than that for females. The strong performance of males can be contrasted with their weaker performance in mathematics in the Junior Certificate examination, and may relate to the different functions of the two assessments, and, perhaps, a greater willingness of males to take risks on an assessment like PISA. The use of a multiple-choice format for some items may also have benefited male students (see Bolger & Kellaghan, 1990). The interaction between the index of books in the home and frequency of absence from school, indicates that, where there are few books in the home, differences in attendance levels (in the two weeks prior to the PISA assessment) are associated with relatively small differences in achievement. However, at higher levels of books in the home, contributions to fitted scores differ considerably, with the lowest scores being achieved by students with high levels of absence (despite the availability of large numbers of books in the home). To the extent that number of books in the home can be interpreted as a proxy for home educational processes (such as support for learning), it can be inferred that absence from school is associated with lower performance, even if the home environment is supportive of learning.

The identification of mathematics self-efficacy and anxiety about mathematics as outcome variables, and their subsequent exclusion from the final mathematics model (even though they explained significant amounts of variance at school and student levels) indicates that care should be exercised in identifying the variables to include in model building, and in interpreting the contributions of variables to achievement. The inclusion of the self-efficacy and anxiety variables would have strengthened the model in terms of its ability to explain variance in achievement (particularly at the student level) at the expense of a clear interpretation of the meaning of the model.

The results of the test-curriculum rating project suggest areas of overlap and divergence between PISA and junior certificate mathematics, and as such they form a useful descriptive tool which could be used as part of a broader consideration of curriculum in Ireland, pointing to possible strengths, weaknesses, and gaps in knowledge and skills.

Comparisons of performance on PISA mathematics and the Junior Certificate examination in mathematics, although complicated by the fact that the purposes of the assessments are very different, nonetheless point to some areas of concern. First, they confirm that the majority of Foundation-level students are at or below the most basic levels of mathematics assessed in PISA. Second, the high proportions of Ordinary-level students scoring at or below Level 1 suggest something of a mismatch, in some cases, between the syllabus level taken and the ability levels of students, which may in turn potentially disguise the numbers of students entering senior cycle who may be in need of additional support.

REFERENCES

Bolger, N., & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement, 27,* 165-174.

Cosgrove, J., Shiel, G., Sofroniou, N., Zastrutzki, S., & Shortt, F. (2005). *Education for life: The achievements of 15-year-olds in Ireland in the second cycle of PISA.* Dublin: Educational Research Centre.

Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). London: Arnold.

Goldstein, H. (2004). International comparisons of student achievement: Some issues arising from the PISA study. *Assessment in Education, 11*, 319-330.

Kaiser, G., Leung, F.K.S., Romberg, T., & Yashenko, I. (2002). International comparisons in mathematics education: An overview. *Proceedings of the International Conference in Mathematics, Beijing, Vol. I,* pp. 631-646. *http://arxiv.org/ S_cache/math/pdf/0212/0212416.pdf*

Longford, N. T. (1993). *Random coefficient models.* Oxford: Oxford University Press.

Lyons, M., Lynch, K., Close, S., Sheerin, E., & Boland, P. (2003). *Inside classrooms: The teaching and learning of mathematics in social context.* Dublin: Institute of Public Administration.

McCullagh, P., & Nelder, J. (1989). *Generalised linear models* (2nd ed.). London: Chapman.

OECD (Organisation for Economic Co-operation and Development). (1999). *Measuring student knowledge and skills: A new framework for assessment.* Paris: Author.

OECD. (2003). *The PISA 2003 assessment framework.* Paris: Author.

OECD. (2004). *Learning for tomorrow's world: First results from PISA 2003.* Paris: Author.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus.* New York: Springer.

Prais, S. J. (2003). Cautions on OECD's recent education survey (PISA). *Oxford Review of Education, 29,* 139-163.

Smithers, A. (2004, May). *England's education: What can be learned by comparing contries? Report to the Sutton Trust.* Liverpool: Centre for Education and Employment Research. *http://www.suttontrust.com/reports/pisa_publication.doc*

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modelin*g. London: Sage.