Education for Life

The Achievements of 15-Year-Olds in Ireland in the Second Cycle of PISA

Judith Cosgrove
Gerry Shiel
Nick Sofroniou
Sarah Zastrutzki
Fionnuala Shortt

Educational Research Centre

Education for Life

The Achievements of 15-Year-Olds in Ireland
in the Second Cycle of PISA

Education for Life

THE ACHIEVEMENTS OF 15-YEAR-OLDS IN IRELAND
IN THE SECOND CYCLE OF PISA

Judith Cosgrove
Gerry Shiel
Nick Sofroniou
Sarah Zastrutzki
Fionnuala Shortt

Educational Research Centre

Cataloguing-in-publication data

Cosgrove, Judith
Education for life: The achievements of 15-year-olds in Ireland in the second cycle of PISA… /
Judith Cosgrove, Gerry Shiel, Nick Sofroniou, Sarah Zastrutzki and Fionnuala Shortt
Dublin: Educational Research Centre

xxiv, 276 p., 30 cm.

ISBN: 0 900440 19 8

Programme for International Student Assessment (Project)
Mathematics (Secondary) – Ireland
Reading (Secondary) – Ireland
Science (Secondary) – Ireland
Problem Solving (Secondary) – Ireland
Academic Achievement
Educational Surveys – Ireland

2005

I Title. II Shiel, Gerry. III. Sofroniou, Nick.

371.262-dc/21

Cover design by eprint Limited, Dublin

Printed in the Republic of Ireland by eprint Limited, Dublin

Contents

List of Tables, Figures and Insets

List of Tables, Figures and Insets (continued)

Tables **Page**

List of Tables, Figures and Insets (continued)

List of Tables, Figures and Insets (continued)

List of Tables, Figures and Insets (continued)

Tables **Page**

List of Tables, Figures and Insets (continued)

List of Tables, Figures and Insets (continued)

List of Tables, Figures and Insets (continued)

List of Tables, Figures and Insets (continued)

Insets

List of Tables, Figures and Insets (continued)

Preface

The Programme for International Student Assessment (PISA) is a collaborative project of the member states of the Organisation for Economic Co-operation and Development (OECD). Its purpose is to address the question of how well young people are equipped for future participation in society, in the labour force, as citizens, and as individuals through an assessment of some key skills. As such, its focus moves away from traditional school-based learning towards a more literacy-based approach that emphasises skills and knowledge needed for future learning and living.

PISA runs in three-yearly cycles. In the first (PISA 2000), reading was the main focus of the assessment, with mathematics and science assuming the status of minor assessment domains. In the second cycle (PISA 2003), the main focus is on mathematics, and there are three minor domains: reading, science, and a new domain – cross-curricular problem solving. In the third cycle (PISA 2006), science will be the major domain, and reading and mathematics will be minor domains. PISA also monitors performance over time. In 2003, it is possible to compare outcomes in mathematics, reading and science with the outcomes for 2000 in countries which participated in both assessment cycles, to give a preliminary indication of changes in achievements over time.

PISA 2000 assessed students in 32 countries/regions (and an additional 11 countries in 2002). In 2003, the number of participating countries is 41, including all 30 OECD member countries and 11 non-OECD (partner) countries. It is expected that 58 countries will participate in PISA 2006.

Several reports based on PISA 2000 have been published by the OECD (see OECD, 2001 and http://www.pisa.oecd.org). A national report for Ireland (*Ready For Life? The Literacy Achievements of Irish 15-Year Olds With Comparative International Data*; Shiel, Cosgrove, Sofroniou, & Kelly, 2001) was published, in full and summary form. A report focusing on reading literacy for teachers and educationalists in Ireland was also published (*A Teacher's Guide to the Reading Literacy Achievements of Irish 15-Year Olds;* Cosgrove, Sofroniou, Kelly, & Shiel, 2003). The present report is the second national publication on PISA 2003 and presents the detailed findings for Irish 15-year-olds in a national context. It was preceded in December 2004 by a summary report which outlined the key findings, and which may be downloaded from http://www.erc.ie/pisa.

This report begins by outlining the key features of the PISA survey design and describes how student skills in the four assessment domains were measured (Chapter 1). Next, the performance of Irish students in earlier international assessments, including PISA 2000, is described, and key variables associated with performance in those assessments are identified (Chapter 2). In Chapter 3, achievement outcomes of students in PISA 2003 mathematics, reading and science are described. In the case of mathematics, achievements are reported both on an overall achievement scale and on subscales representing four mathematical content areas (Space & Shape, Change & Relationships, Quantity, and Uncertainty). Outcomes for 2000 and 2003 are compared for mathematics, reading and science. Chapter 4 examines some key student-level and school-level variables associated with student achievement in Ireland. Chapter 5 builds on Chapter 4 by examining the associations of a number of school- and student-level variables simultaneously, through multilevel models of achievement in mathematics, reading and science. Chapter 6 explores outcomes on PISA in relation to the Irish Junior Certificate syllabus and examinations, with a particular focus on the content and assessment of mathematics. Chapter 7 describes student performance on the assessment of cross-curricular problem solving. Finally, Chapter 8 summarises findings and draws a number of conclusions and implications. The report is designed in such a way that chapters may be read as stand-alone documents; where appropriate, cross-references to other parts of the report are made.

To illustrate the nature of the test items encountered by students in PISA 2003, an appendix to the report contains some sample assessment tasks. Additional sample tasks can be found on http://www.erc.ie/pisa. Sample items also feature in the initial OECD reports for PISA 2003 (OECD, 2004b, c) and the PISA 2003 assessment framework document (OECD, 2003).

Abbreviations and acronyms used in this report are listed on page xvii.

We would like to acknowledge the contributions of a number of individuals during the implementation of PISA 2003 in Ireland and the production of this report. Firstly, thanks are due to the 145 schools and 3880 students in Ireland that took part in the assessment in March, 2003. Without their support and assistance, particularly the assistance of the School Contact in each school, the study could not have been carried out. Thanks also to the 27 schools and 734 students who participated in the pilot phase of this survey in March 2002. Outcomes of the pilot were invaluable in the selection of the final test and questionnaire items, and making refinements to survey procedures.

Thanks are extended to the PISA consortium and OECD Secretariat, particularly Andreas Schleicher, Claudia Tamassia, and Christian Monseur, for assistance and advice in the course of analysing the data for this report. We would also like to acknowledge the advice of Murray Aitkin (University of Newcastle) on nonparametric maximum likelihood estimation, which is applied to the comparison of country mean scores in Chapter 3.

Thanks are due to Doreen McMorris (Department of Education and Science), Chair of the PISA national advisory committee, for continued support for our work on PISA, and advice on the content of this report. Thanks also to the other PISA national advisory committee members for their detailed and insightful comments and feedback on drafts of this report: Declan Kennedy (University College, Cork), Bill Lynch (National Council for Curriculum and Assessment), Tom Mullins (University College, Cork), and Elizabeth Oldham (University of Dublin, Trinity College). Thanks to Seán Close (St Patrick’s College, Dublin), who recently joined the committee and offered detailed advice on the content of the report, particularly regarding the interpretation of proficiency levels and student performance in mathematics.

We wish to thank the individuals involved in the rating of the PISA mathematics tasks with respect to the Junior Certificate mathematics syllabus and examinations: John Evans, Elizabeth Oldham, and Peter Tiernan.

Thanks are due to our colleagues in the Educational Research Centre, Mary Rohan, John Coyle, and Stephen McMahon for their administrative and technical support, and to Amy Kelly, who contributed to the project as a Research Assistant until May 2004.

Throughout this report, the terms *mathematics*, *reading* and *science* are used to denote the PISA domains of mathematical literacy, reading literacy and scientific literacy. This is consistent with the approach adopted in the international report on PISA (OECD, 2004b).

List of Abbreviations and Acronyms

BRR	balanced repeated replication (design)
CI95L	lower limit of 95% confidence interval
CI95U	upper limit of 95% confidence interval
df	degrees of freedom
ESCS	economic, social and cultural status scale
IAEP	International Assessment of Educational Progress
IALS	International Adult Literacy Survey
ICC	intra-cluster correlation
IEA	International Association for the Evaluation of Educational Achievement
INES	International Indicators of Education Systems
IRT	item response theory
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ISEI	International Socio-Economic Index
JCE	Junior Certificate Examination
NCCA	National Council for Curriculum and Assessment
OECD	Organisation for Economic Co-operation and Development
PISA	Programme for International Student Assessment
RME	Realistic Mathematics Education
SD	standard deviation
SE	standard error
SED	standard error of the difference
SES	socioeconomic status
SIMS	Second International Mathematics Study
TIMSS	Third International Mathematics and Science Study

Executive Summary

The second cycle of the OECD Programme for International Student Assessment (PISA 2003) was implemented in 30 OECD member countries (including Ireland) and in 11 non-OECD (partner) countries in Spring/Autumn 2003. In Ireland, a nationally representative sample of 3880 15-year-olds in 145 post-primary schools participated. Tests consisting of items from one major assessment domain (mathematics) and three minor assessment domains (reading literacy, science, and cross-curricular problem solving) were administered to students in their schools.[1] Students also completed a questionnaire which sought information on their attitudes, interests and home backgrounds. Their principal teachers completed a questionnaire on issues such as school management and staffing. These questionnaires provided contextual information that facilitated the interpretation of student performance. The first international reports on PISA 2003 were released by the OECD in December, 2004 (OECD, 2004b, c). The Educational Research Centre published a summary version of the current report at the same time (Cosgrove, Shiel, Sofroniou, Zastrutzki, & Shortt, 2004).

PISA 2003: KEY FINDINGS FOR IRELAND

Achievement in Mathematics, Reading, Science and Problem Solving

Students in Ireland achieved mean scores that were significantly higher than the corresponding OECD country average scores in reading literacy and science. In mathematics and cross-curricular problem solving, Ireland's mean scores did not differ from the corresponding OECD averages.

On the combined (overall) mathematics scale Ireland ranked 20th of 40 participating countries (95% confidence interval [CI] for Ireland's ranking = 16th to 21st), and 17th of 29 OECD countries[2] (95% CI = 14th to 19th). Ireland's mean score in the reading literacy scale gave it a rank of 7th of 40 countries (95% CI = 5th to 11th), and 6th of 29 OECD countries (95% CI = 5th to 9th). Ireland ranked 16th of 40 countries (95% CI = 12th to 20th) and 13th of 29 OECD countries (95% CI = 9th to 16th) on the science scale. In cross-curricular problem solving, Ireland ranked 21st of 40 countries (95% CI = 20th to 22nd) and 18th of 29 OECD countries (95% CI = 17th to 19th).

The highest-performing countries in PISA 2003 included Hong Kong-China (ranked 1st in mathematics, 10th in reading, 3rd in science and 2nd in problem solving), Korea (3rd in mathematics, 2nd in reading, 4th in science, and 1st in problem solving), and Finland (2nd in mathematics, 1st in reading, 1st in science, and 3rd in problem solving).

Classification of Countries into Performance Groupings

Using a non-parametric maximum likelihood estimation technique, the performance of countries was classified into empirically distinct groupings. In mathematics, it was possible to distinguish 6 performance groups (Ireland is in the second highest grouping); 10 in reading (Ireland lies in the third highest); 7 in science (Ireland is in the third highest); and 6 in problem solving (Ireland lies between the second and third highest groups).

[1] While all students took at least one block of mathematics items, item blocks in the minor domains were systematically rotated across booklets.

[2] The results for 29 rather than 30 OECD countries are reported since the United Kingdom did not meet the sampling standards required to ensure reliable achievement estimates.

Achievement on the Mathematics Subscales

In addition to combined mathematics, PISA 2003 reported performance on four mathematical subscales (Space & Shape, Change & Relationships, Quantity, and Uncertainty). Although several countries showed consistently high performance on all four subscales (e.g., Finland, Korea, the Netherlands), Ireland's performance showed some variation. Irish students achieved mean scores that were significantly above the OECD country average on two subscales – Uncertainty (one-sixth of a standard deviation higher), and Change & Relationships (one-twentieth of a standard deviation higher). Performance did not differ significantly from the OECD average on the Quantity subscale, and was significantly below it on the Space & Shape subscale (by one-fifth of a standard deviation).

Distribution of Achievement in Mathematics, Reading, Science and Problem Solving

An examination of the percentages of students scoring at each of six proficiency levels indicates that the performance of Irish students on the overall mathematics scale was characterised by comparatively few very high and very low achievers. For example, 16.8% scored at or below Level 1 (the lowest proficiency level assessed), compared to just over 21% of students on average across the OECD. Just over 11% of Irish students scored at Levels 5 and 6 (the highest two levels), compared to an OECD average of 14.6%.

A similar pattern is evident for problem solving. Just 12.3% of Irish students achieved Level 3 (the highest level) in comparison with an OECD country average of 18.2%, and just 12.5% achieved scores below Level 1 (the lowest level), in comparison with the OECD country average of 17.3%.

Ireland's overall high average performance in reading was characterised by comparatively high achievement at the lower end of the distribution and moderate achievement at the upper end. The percentage of Irish students who scored at or below the lowest reading proficiency level, Level 1 (11.0%) is well below the corresponding OECD average (19.1%). On the other hand, the percentage of students in Ireland with reading scores at the highest level, Level 5 (9.3%) does not differ significantly from the OECD country average (8.3%).

Proficiency levels were not developed for science. However, a comparison of scores at key percentile points indicates a pattern similar to mathematics and problem solving. The science score at the 10th percentile for Ireland is 22.3 points higher than the OECD average at this point (361.6), while the Irish score at the 90th percentile is 9.7 points lower than the OECD average (634.2).

The mean PISA mathematics scores differed for students who took the Junior Certificate mathematics examination at Higher, Ordinary, and Foundation levels (in the region of one standard deviation between Foundation and Ordinary levels, and one standard deviation between Ordinary and Higher levels). A similar pattern was found for Junior Certificate English for scores on the PISA reading literacy test. Students taking Junior Certificate science at Higher and Ordinary levels also differed significantly in their scores on the PISA science test (by about two-thirds of a standard deviation).

Achievement Differences Between Students

In Ireland, the standard deviations (measures of dispersion around mean scores) in all four PISA 2003 domains and on the mathematics subdomains are smaller than in most other countries. For example, Ireland has the lowest standard deviation among OECD countries on the problem solving test, although those for Finland and Iceland are only marginally larger. These narrow variations of achievement in Ireland can be interpreted as indicating greater homogeneity in achievement outcomes compared to countries in which the dispersions of scores are wider.

Achievement Differences Between Schools

In Ireland, between-school variation in mathematics (corresponding to the intra-cluster correlation, ICC) is 16.7% which is below the OECD country average (32.7%). The Irish ICCs for reading, science, and problem solving are 22.5%, 16.2% and 15.7% respectively. These are also lower than the corresponding OECD country averages (31.4%, 29.9% and 31.6%). Thus, relatively little of the variation in achievement in any of the domains is between schools in Ireland; most is within schools (between classes and students).

Gender Differences in Achievement

Males significantly outperformed females in Ireland on the combined mathematics scale (by 14.8 points; about one-sixth of a standard deviation), and females significantly outperformed males on the reading literacy scale (by 29.0 points; about one-third of a standard deviation). The mean scores of male and female students in Ireland on the science and problem-solving scales did not differ significantly. On the mathematics subscales, the largest gender difference was associated with Shape & Space (25.5 points or close to one-third of a standard deviation in favour of males). The pattern of gender differences in the PISA 2003 domains and subdomains is consistent with those found in most participating countries. It is also consistent with gender differences in PISA 2000.

Student-Level Variables and Achievement

Student-level variables tended to show similar associations with achievement in all four domains in Ireland. These include student socioeconomic status, lone parent status, number of siblings, frequency of absence from school, availability of educational resources in the home (e.g., a quiet place to study), number of books in the home, and time spent on homework/study. Two additional variables, self-efficacy in mathematics and attitude towards mathematics, are associated with performance in mathematics. The study of science as a school subject for the Junior Certificate is associated not just with performance in science, but also with performance in the other three domains.

School-Level Variables and Achievement

School-level variables associated with differences in achievement in Ireland include school size, school sector (whether secondary, vocational or community/comprehensive), school designated disadvantaged status, school gender composition, Junior Certificate Examination fee waiver entitlement, disciplinary climate in mathematics classes, ratio of computers to students, and instructional time at school and in mathematics classes.

Explaining Achievement

The performance of Irish students in mathematics, reading literacy, and science was examined within a multilevel modelling framework, which allows the effects of a number of variables at both school and student levels to be examined simultaneously. The models for the three domains revealed broadly similar results. All three models indicated that two school-level variables in particular were associated with achievement: school socioeconomic status (based on the percentage of students in the school in receipt of a fee-waiver for the Junior Certificate Examination), and school disciplinary climate (a measure of the learning environment in mathematics classes). School sector (whether a school was secondary, community/comprehensive, or vocational) was not required in the final models, as differences related to it were explained by other variables.

The combination of student-level variables that explained a substantial portion of the variation in performance were gender, frequency of absence from school, lone-parent status, number of siblings, number of books in the home, access to home educational resources, and current grade level. An additional student variable, study of science at school, was included in the final model for science. In mathematics, the final model explained 78.8% of between-school variation, and 29.6% of the variation within schools. The corresponding figures for reading were 81.4% and 35.4%, and for science, 80.2% and 31.2%.

Students' perceived self-efficacy in mathematics and anxiety about mathematics explained additional variation in achievement in mathematics over and above the other school and student variables in the model. However, conceptual and theoretical difficulties occur with using these as explanatory variables; they may be better considered as educational outcomes conjointly with achievement outcomes.

Comparison of PISA Mathematics and the Irish Mathematics Curriculum

A comparison of the Junior Certificate mathematics syllabus/examination with the PISA mathematics assessment instrument revealed considerable differences between the two, both in terms of the mathematical concepts that were assessed and the ways in which they were presented (particularly the contexts in which items were embedded). On the basis of ratings of PISA mathematics items made by trained raters with reference to Junior Certificate mathematics syllabus documents, one would expect Third year students to be unfamiliar with the mathematical concepts underlying between 30% and 50% of PISA items (depending on the syllabus level studied/taken), and with the contexts in which 66% to 80% of items were embedded.

When each PISA mathematics item was classified in terms of the Junior Certificate mathematics topic area in which its underlying concept was expected to be taught, it was found that between 29% and 49% of the items were not on the Junior Certificate mathematics syllabus (depending on the syllabus level considered). Of PISA items whose concepts could be located in the syllabus, a majority were in the topic areas of applied arithmetic and measure, and statistics. Few or none of the PISA items were located in the Junior Certificate mathematics topic areas of functions and graphs, sets, geometry, and trigonometry.

Relationships Between PISA Domains and Performance on the Junior Certificate

Correlations between student performance on PISA mathematics/Junior Certificate mathematics, PISA reading literacy/Junior Certificate English, and PISA science/Junior Certificate science range from .68 to .75, indicating an appreciable degree of association. Correlations amongst Irish student scores in the four PISA domains, however, are higher, ranging from .80 to .90. The correlation between mathematics and problem solving (.90) is particularly high, and noteworthy, since it was intended that the assessment of problem solving would tap skills independent of those assessed in mathematics.

Comparing Achievement in 2000 and 2003

OECD average scores did not change in reading or science between 2000 and 2003. In mathematics, where it was possible to compare differences only on the Change & Relationships and Space & Shape subscales, the OECD average score on the former increased between 2000 and 2003, while the OECD average score on the latter remained unchanged. Average performance in Ireland on the two mathematics subscales did not change between 2000 and 2003; nor did performance on the science scale. However, average performance in reading was significantly lower in 2003 than in 2000, albeit by just one-tenth of a standard deviation. Irish performance in reading at the 75th, 90th, and 95th percentiles also decreased significantly.

SOME IMPLICATIONS

Overall performance in mathematics. The overall mean score of students in Ireland, although not significantly different from the OECD country average, is lower than the means of students in a number of countries. This prompts the question: are current standards in mathematics adequate, and do they meet the current and future needs of students in Ireland? In addressing these questions, it should be recognised that there are substantial differences between PISA mathematics and the Junior Certificate mathematics syllabus in their underlying philosophies and approaches (e.g., the framework for PISA mathematics is grounded in the Realistic Mathematics Education movement, while Junior Certificate mathematics is not).

Performance of high achievers in mathematics. The relatively low performance of higher-achieving students in mathematics in Ireland is noteworthy and suggests that any forthcoming review of mathematics at post-primary level should consider this finding, with a view to identifying ways in which performance of high achievers might be enhanced.

Gender differences in mathematics. Male students in Ireland achieved a mean score on combined mathematics that is significantly higher than that of females. Gender differences favouring males were also observed in the four mathematics subdomains, with the largest occurring on the Space & Shape scale. These differences, although consistent with the general pattern observed across most participating countries, contrast with those for students taking the Junior Certificate mathematics examination in 2003, in which females outperformed males by about one-half of a grade (one-sixth of a standard deviation). Further investigation of these differences is merited.

Concepts underlying PISA and Junior Certificate Mathematics. The weak match between PISA mathematics and the Junior Certificate mathematics syllabus suggests that any future review of mathematics education at post-primary level should consider if important mathematical content is absent from the syllabus. However, any debate around the differences between PISA and Junior Certificate mathematics will need to take into account that some topic areas of the Junior Certificate mathematics syllabus are not assessed by PISA (e.g., sets, geometry and trigonometry), that some PISA concepts (e.g., probability) only appear on mathematics syllabi at Senior Cycle level, and that students may acquire mathematical concepts outside mathematics classes.

Between-school variation in achievement. In Ireland, relatively small proportions of the variation in achievement are attributable to differences between schools. This indicates that, relative to many other countries, Irish schools are more similar to one another in terms of achievement. This may be partly due to the relatively homogeneous cultural composition of the school-going population in Ireland. Any efforts to increase performance in mathematics and other areas would need to ensure that differences between schools do not show a concurrent increase.

Performance in reading in 2000 and 2003. Reading is the only PISA domain in which a difference in the performance of students in Ireland was observed between 2000 and 2003. The mean score was significantly lower in 2003 than in 2000, while the scores of students at the 75th, 90th and 95th percentile points in Ireland were also significantly lower. It is recommended that caution be exercised in any interpretation of change until additional data, gathered over a longer period of time, become available. Despite the decline, Ireland remains among the highest performers on PISA reading literacy.

Study of science. As in PISA 2000, students in PISA 2003 who reported that they did not study science as a subject for the Junior Certificate Examination (5.2% of males; 14.6% of females; 9.9% of all students) achieved a mean score that was lower by 69.7 points (four-fifths of a standard deviation) than that of students who took science as a subject. Students who did not take science also performed less well in mathematics and reading than those who did. It would seem important that all First-year students are well informed about the potential benefits of choosing science as a subject (e.g., by providing short 'taster' courses in science). It would also seem important to make efforts to develop the scientific knowledge of all Junior Cycle students, although it may not be necessary for all of them to take science in the Junior Certificate Examination to accomplish this.

School socioeconomic status and performance. Multilevel models of achievement in mathematics, reading, and science highlight the contributions of individual- and school-level socioeconomic status to achievement. The level of disadvantage associated with the school that a student attends (based on the percentage of students who were entitled to a Junior Certificate Examination fee waiver) was significantly associated with achievement in combined mathematics, reading, and science, even when adjustments had been made for other school and student variables, including student socioeconomic status. This finding justifies current efforts to target resources on schools with large numbers of students from disadvantaged backgrounds, although the effects of such efforts need to be carefully monitored.

School disciplinary climate and performance. After adjusting for the effects of other school- and student-level variables, students in schools with a high positive average disciplinary climate in their mathematics class (as perceived by the students themselves) had higher expected mean scores in all assessment domains than students in schools with medium and low levels. Further, the correlation between school-level socioeconomic status and disciplinary climate is very low, indicating that school disciplinary climate is weakly associated with school socioeconomic status. These findings suggest that the measurement and nature of disciplinary climate merit further examination.

Absenteeism and performance. Although data on student attendance was limited to the two weeks prior to the PISA assessment (and reasons for these absences are unknown), effects on achievement in reading and science are statistically significant. In the multilevel model of mathematics, there was an interaction between attendance and the index of books in the home whereby students who were absent for three or more days, and who had few books in the home, had the lowest fitted scores. These findings underline the value of supporting the regular attendance of students, especially those from backgrounds where literacy activities may not be emphasised.

1

The PISA 2003 Assessment: Overview and Implementation

The Programme for International Student Assessment (PISA) is an assessment of the knowledge and skills of 15-year-olds that takes place at three-year intervals, under the auspices of the Paris-based Organisation for Economic Co-operation and Development (OECD). The first cycle of PISA (called PISA 2000) was implemented in 2000. The main assessment domain was reading literacy, with mathematics and science as minor domains. Initial international results of PISA 2000 were published in *Knowledge and Skills for Life: First Results of PISA 2000* (OECD, 2001a). An Irish national report, *Ready for Life: The Literacy Achievements of Irish 15-year-olds* (Shiel, Cosgrove, Sofroniou, & Kelly, 2001), provided a detailed analysis of the performance of Irish students in the study. The focus of these and subsequent reports on PISA 2000 (e.g., Kirsch, de Jong, Lafontaine, McQueen, Mendelovits & Monseur, 2002; Sofroniou, Shiel, & Cosgrove, 2002) was on explaining the outcomes of PISA 2000 with reference to the characteristics of schools, students, and their families in ways that are relevant to policy makers in participating countries. The issue of equity in learning outcomes, both within and across countries, underpinned much of this work.

Mathematics was the major assessment domain in PISA 2003, while reading literacy, science, and a new domain, cross-curricular problem solving, were minor domains. Students enrolled in school- and work-based educational programmes in 41 countries, including all 30 OECD member countries (Inset 1.1), sat paper-and-pencil tests in these domains, and also completed a Student Questionnaire that sought information about students' home background, attitudes to mathematics, learning strategies, educational and occupational aspirations, and familiarity with information and communication technologies. School principals completed a School Questionnaire that asked about school admission policies, resource availability, management, assessment policies and teacher qualifications. In Ireland, the mathematics teachers of participating students completed a Teacher Questionnaire that asked about instructional practices, availability of resources for teaching mathematics, and implementation of the revised Junior Certificate Mathematics Syllabus (Department of Education and Science, 2000). The findings of the Teacher Questionnaire will be addressed in a later publication.

PISA is not intended to provide a direct measure of students’ mastery of specific curricular content. Rather, key content and learning processes that are deemed important for the present and future lives of 15-year-olds as individuals and as members of society have been identified by panels of international domain specialists, and are embedded in ‘real life’ problems. This ‘literacy-based’ perspective is reflected in the full titles of the assessment domains (reading literacy, mathematical literacy, and scientific literacy), in the definition of each domain (including problem solving), and in the tasks that students were asked to attempt.

Inset 1.1. Countries Participating in PISA 2003

OECD Countries			*Partner Countries**
Australia	Iceland	Portugal	Brazil
Austria	Ireland	Slovak Republic**	Macao-China**
Belgium	Italy	Spain	Hong Kong-China
Canada	Japan	Sweden	Indonesia
Czech Republic	Korea (Rep. of)	Switzerland	Latvia
Denmark	Luxembourg	Turkey**	Liechtenstein
Finland	Mexico	United Kingdom	Russian Federation
France	Netherlands	United States	Serbia
Germany	New Zealand		Thailand
Greece	Norway		Tunisia
Hungary	Poland		Uruguay

*Partner countries are not OECD member countries **New to PISA in 2003

The distinction that PISA makes between major and minor domains (Inset 1.2) is reflected in the proportion of test items allocated to each domain. In PISA 2003, across all 13 test booklets, there were 85 mathematics items, 28 reading items, 35 science items, and 19 problem-solving items. The designation of one major domain in each assessment cycle means that performance in that domain can be described in detail, often in terms of subdomains, while performance in the minor domains can be described in broader terms. PISA also provides data on changes in achievement over time. The current report looks at differences in performance in reading literacy, mathematics, and science between 2000 and 2003.

The inclusion of problem solving as a minor assessment domain in PISA 2003 (see Inset 1.2) reflects an interest in measuring knowledge and skills that can be applied by students across different subject domains. An important issue that can be examined in PISA 2003 is the extent to which cross-curricular problem solving (as defined and assessed in PISA) is related to, or independent of, problem solving in other assessment domains such as mathematics and science.

Inset 1.2. Content Domains Across Planned Cycles of PISA (2000-2006)

Year	*Major Domain*	*Minor Domains*	*Additional Areas of Interest**
2000	Reading (141)**	Mathematics (32) Science (35)	Equity and literacy; reading attitudes and habits; students' self-regulated learning
2003	Mathematics (85)	Science (35) Reading (28) Cross-Curricular Problem Solving (19)	Variables associated with performance in mathematics; attitudes to mathematics; educational pathways
2006	Science	Reading Mathematics	Information and communication technologies (ICTs); attitudes to science

* These areas are addressed through the administration of questionnaire items.

** Numbers of items for each domain in 2000 and 2003 are indicated in brackets.

PISA stems from, and now feeds into, the work of the OECD INES (Indicators of Education Systems) Project, which gathers cross-nationally comparable indicators on education systems which are published annually in *Education at a Glance* (e.g., see OECD, 2004a). Indicators from PISA include achievement outcomes as well as indicators demonstrating links between background variables and achievement (e.g., gender, interest in mathematics, use of self-regulated learning strategies). These complement other indicators gathered by the INES network, including the financial and human resources invested in education, access to education, and the learning environment in schools.

PISA can also be viewed in the context of current interest in human capital. The OECD (1998) defines human capital as 'the knowledge, skills, competencies and other attributes that are embodied in individuals that are relevant to personal, social and economic well-being' (p. 9). To the extent that the PISA assessment domains measure the knowledge and skills required for future adult life, performance on these domains can be interpreted as indicators of human capital and of the preparedness of students for life-long learning.

According to the OECD (2004b), PISA can be used by countries to: gauge the literacy skills of their students in comparison with students of other participating countries; establish benchmarks for educational improvement, in terms of the performance of other countries, or their capacity to provide high levels of equity in educational outcomes and opportunities; and understand relative strengths and weaknesses of educational systems.

Key features of the PISA 2003 assessment are summarised in Inset 1.3.

Inset 1.3. Key Features of the PISA 2003 Assessment

- An internationally standardised assessment of 15-year-olds, jointly developed by participating countries and administered to over 250,000 students in 41 countries
- A focus on how young people near the end of compulsory schooling can use their knowledge and skills to meet real-life challenges
- An emphasis on the mastery of processes, the understanding of concepts and the ability to function in various situations within each assessment domain
- The administration of paper-and-pencil assessments involving both multiple-choice items and items requiring students to construct their own answers
- The development of a profile of skills and knowledge among students at or near the end of compulsory schooling
- The development of contextual indicators relating results to student and school characteristics
- The development of trend indicators that can track changes over time

The remainder of this chapter is divided into four parts. In the first, the assessment of literacy in PISA 2003 is described with reference to the types of items used and the frameworks for mathematics, reading literacy, science, and problem solving. In the second part, the PISA Student and School Questionnaires are described. The third part describes the implementation of PISA in Ireland focusing on such matters as sampling schools and students and monitoring the implementation of assessment procedures. The fourth part describes some of the procedures used to analyse the PISA data in Ireland.

FRAMEWORK FOR MATHEMATICAL LITERACY (MATHEMATICS)

Mathematics in PISA 2003 is concerned with 'the capacities of students to analyse, reason, and communicate ideas effectively as they pose, formulate, solve and interpret mathematical problems in a variety of situations' (OECD, 2003, p. 24). The focus is on 'real-world' problems, though mathematical problems of the traditional kind, which are not embedded in realistic situations, are also included. Students are expected to draw on their quantitative or spatial reasoning, or other mathematical competencies, to help clarify, formulate or solve problems. In doing so, they are expected to 'make decisions about what knowledge may be relevant, and how it might usefully be applied' (p. 24). The contexts in which students are expected to interpret and solve mathematical problems range from purely mathematical ones to contexts in which the mathematics is embedded and requires mathematisation, i.e. isolating and solving the mathematics problem. Mathematical literacy or mathematics in PISA 2003 is defined as:

> an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgements and to engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen. (OECD, 2003, p. 24)

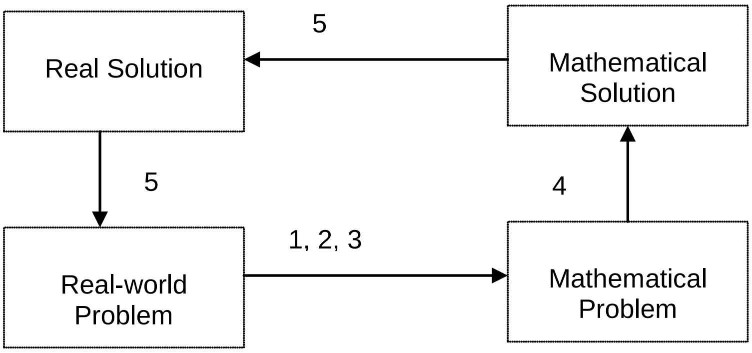
The term 'mathematical literacy' is used to emphasise mathematical knowledge put to functional use in a variety of ways. PISA mathematics presupposes technical aspects of mathematical knowledge, including 'terminology, facts and procedures, as well as skills in performing certain operations and carrying out certain methods' (OECD, 2003, p. 25). It required all these, but also the ability to 'creatively combine these elements in response to the demands imposed by the external situation' (p. 25).

The definition and accompanying framework are heavily influenced by the realistic mathematics education (RME) movement, which stresses the importance of solving mathematical problems in real-world settings (e.g., Freudenthal, 1973, 1981), and by ideas related to situated cognition, which focus on the importance of the activities and situations in which learning is embedded as well as the cognitive dimensions (e.g., Brown, Collins & Duguid, 1989). Central to these, and to the PISA concept of mathematics, is the notion of mathematising (sometimes referred to as 'mathematical modelling') – a process in which mathematicians engage, but which is also deemed to be relevant to the aspects of everyday life that call for the application of mathematical knowledge. Mathematising involves:

1. starting with a problem situated in reality;
2. organising the problem according to mathematical concepts;
3. gradually 'trimming away the reality' through such processes as making assumptions about which features of the problem are important, generalising and formalising the problem;
4. solving the mathematical problem; and
5. making sense of the mathematical solution in terms of the real situation.

The process of mathematising, which includes both horizontal and vertical dimensions, is illustrated in Figure 1.1 (numbers indicate the dimensions of mathematisation described above). The horizontal dimension refers to students discovering the mathematical tools which can help them to organise and solve a problem located in a real-life situation (Steps 1-3 in Figure 1.1). The vertical dimension entails a process of reorganisation within the confinements of mathematical concepts, and includes generalising, refining algorithmic models, and using different modes to reach a solution (Treffers, 1987).

Figure 1.1. The Mathematisation Cycle

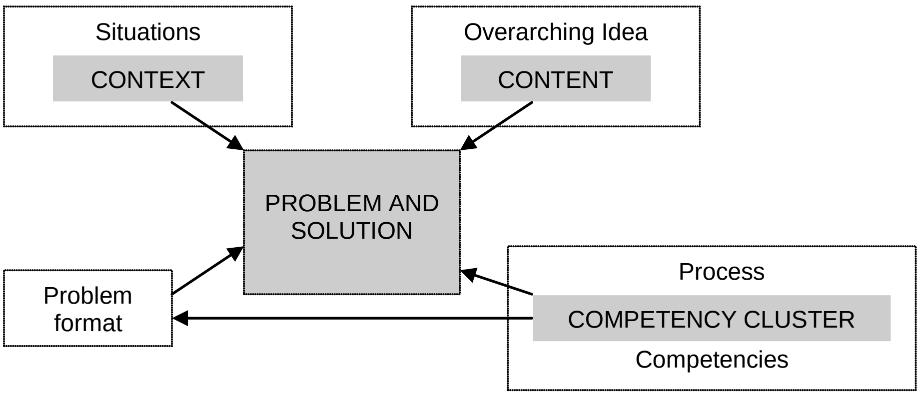


Source: OECD (2003), Figure 1.3, p. 38

Mathematics processes represented by numbers 1-5 are described on this and the previous page.

In an ideal situation, PISA would assess students' abilities to use mathematics to solve real-world problems in real-life settings. Students would work in groups, each one contributing ideas on how a problem might be solved, and what the results mean. However, given the constraints under which PISA operates, it was only possible to administer paper-and-pencil tests to students, on an individual basis, and in a limited time frame. The remainder of this section describes how this was accomplished. Three dimensions of the framework are discussed: mathematical aspects or 'overarching ideas', mathematical competencies, and situations and contexts (see Figure 1.2).

Figure 1.2. The Components of Mathematics in PISA 2003



Source: OECD (2003), Figure 1.2, p. 30.

Mathematical Content – The Four 'Overarching Ideas'

PISA identifies four 'overarching' ideas: Space & Shape, Change & Relationships, Quantity, and Uncertainty. It is argued that these ideas 'meet the requirements of historical development [of mathematics], coverage of domain, and reflection of the major threads of school curricula' (OECD, 2003, p. 35). Moreover, it is acknowledged that these are not mutually exclusive, but can be viewed as 'generalised content dimensions' that include elements of mathematics curricula, such as arithmetic, measurement, geometry, algebra, and probability and statistics.

Space and Shape

Items tapping this subdomain are concerned with recognising and understanding geometric patterns and identifying such patterns in abstract and real-life representations. They entail identifying similarities and differences among shapes, and recognising shapes in different representations and dimensions. They involve understanding the properties of objects and their relative positions, how three-dimensional objects can be represented in two dimensions, and what perspective is and how it functions. Examples of Space & Shape tasks in Appendix A are Staircase, Number Cubes, and Carpenter.

Change and Relationships

Tasks in this subdomain require students to analyse data to determine the kind of relationship that is present. Relationships are often represented using equations or inequalities, but relations of a more general nature, such as equivalence, divisibility, and inclusion, are also relevant. Tasks involve identifying mathematical manifestations of change, as well as functional relationships, and dependency among variables.

This subdomain also stresses functional thinking – thinking in terms of and about relationships in a variety of forms including symbolic, algebraic, graphical, tabular and geometric. As different representations may serve a variety of purposes, and have different properties, translation between representations may be of key importance in dealing with this area of mathematics. Examples of Change & Relationships tasks in Appendix A are Growing Up and Walking.

Quantity

Questions classified under Quantity assess students' ability to quantify, to understand relative size, and to recognise numerical patterns and use numbers to represent quantities and quantifiable attributes of real-world objects (counts and measures). Quantity also deals with the processing and understanding of numbers that are represented in various ways.

An important aspect of dealing with Quantity is quantitative reasoning. This involves number sense, the ability to represent numbers in a variety of ways, mental arithmetic, estimation, and understanding the meaning of operations and the magnitude of numbers. Examples of Quantity tasks in Appendix A are Exchange Rate and Skateboard.

Uncertainty

Uncertainty is viewed as comprising two factors: data and chance. These correspond, respectively, to statistics and probability in mathematics curricula. Specific activities in this subdomain are collecting data to test hypotheses and inform decision-making, data analysis and display/visualisation, probability, and inference. Examples of Uncertainty tasks in Appendix A are Exports and Test Scores.

Mathematical Competencies and Competency Clusters

PISA places a strong emphasis on the ability to solve mathematical problems situated in real-life contexts. As described earlier, this requires students to use the mathematical skills and competencies they have acquired in the process of mathematisation. PISA identifies a range of competencies, which, when taken together, are viewed as constituting ‘comprehensive mathematical competence’. Moreover, it is recognised that each competency can be possessed at different levels of mastery (which may or may not match the level of mastery required by specific PISA mathematical problems). PISA identifies eight mathematical competencies: thinking and reasoning; argumentation; communication; modelling; problem posing and solving; representation; using symbolic, formal and technical language and operations; and using aids and tools (for example, calculators).

PISA does not develop test items around each of the eight competencies. Rather, elements of each are embodied in three broad competency clusters: the Reproduction cluster, the Connections cluster, and the Reflection cluster. Each mathematical problem is classified according to the cluster that best describes the competencies that the student is required to draw on to solve the problem. The competency clusters are assumed to form a hierarchy, with the Connections cluster building on the skills from the Reproduction cluster, and the Reflection cluster building on the skills from the Connections cluster. While each cluster is intended to feature both easy and difficult items, in practice there tends to be a hierarchy of difficulty. Key features of each competency cluster are described below (see Table 1.1 for a summary).

Table 1.1. The PISA 2003 Mathematics Competency Clusters

Reproduction Cluster	*Connections Cluster*	*Reflection Cluster*
Reproducing representations, definitions and facts	Integrating and connecting across content, situations and representations	Complex problem solving and posing
Interpreting simple, familiar representations	Non-routine problem solving, translation	Reflecting on, and gaining insight into, mathematics
Performing routine computations and procedures	Interpretation of problem situations and mathematical statements	Constructing original mathematical approaches
Solving routine problems	Using multiple well-defined methods	Communicating complex arguments and complex reasoning
	Engaging in simple mathematical reasoning	Using multiple complex methods
		Making generalisations

Source: Adapted from OECD (2003), Figure 1.4, p. 49

The Reproduction Cluster

Competencies required in this cluster involve reproduction of practised or familiar skills. They include: understanding and handling mathematical concepts in the contexts in which they were first introduced or practised; recognising and reproducing standard pure and applied problems in closed form; solving problems using standard approaches and routines; interpreting familiar practised representations of well-known mathematical objects; and handling simple statements and expressions containing symbols and formulae. Examples of Reproduction items include:

- solving an equation such as $7x - 3 = 13x + 15$;
- expressing a percentage as a fraction;
- calculating the perimeter of a circle; and
- calculating the simple interest on a sum of money.

Examples of items in the Reproduction cluster in Appendix A include Exchange Rate Q1 and Q2, Staircase Q1, Exports Q1, Growing Up Q1 and Q2, Skateboard Q1 and Q2, and Walking Q1.

The Connections Cluster

Items in this cluster involve integrating and connecting material from the various overarching ideas or from different mathematical curriculum strands, or linking different representations of a problem. The competencies include: distinguishing between definitions and assertions; engaging in simple mathematical reasoning without distinguishing between proofs and broader forms of argument and reasoning; explaining computations and their results (usually in more ways than one); translating 'reality' into mathematical structures; choosing and switching between different forms of representation of mathematical objects and situations; decoding and interpreting basic symbolic and formal language in less well-known contexts and situations; and using familiar aids and tools in contexts that are different from those in which their use was introduced and practised. Examples of 'Connections' items include:

- Given the distance two children live from the school, finding out how far they live from each other; and
- Given the diameter and cost of two pizzas, finding which is the better value for money, and showing one's reasoning.

Examples of items in Appendix A that in the Connections cluster include Exports Q2, Growing Up Q3, Skateboard Q3, Number Cubes Q2, Walking Q3, Test Scores Q1, Carpenter Q1, Robberies Q1, and Internet Relay Chat Q1.

The Reflection Cluster

This cluster builds on the skills from the Connections cluster but also requires a stronger element of reflectiveness from students. The cluster includes advanced reasoning, argumentation, abstraction, generalisation, and modelling applied to new contexts. In addition to competencies in the Reproduction and Connections clusters, it requires: distinguishing between definitions, theorems, conjectures, hypotheses, and assertions about special cases; assessing and constructing chains of mathematical arguments of different types; explaining matters that include complex, sometimes logical, relationships; translating 'reality' into mathematical structures that may differ considerably from what students are familiar with; monitoring the modelling process, and validating the resulting model; solving problems by invoking and using more original problem-solving processes in which connections are made between different mathematical areas and modes of representation and communication (schemata, tables, graphs, words, and pictures); interpreting less familiar representations of mathematical objects, and distinguishing between different forms of representation; dealing with complex statements and expressions and with unfamiliar symbolic or formal language; and using aids and tools in contexts that are quite different from those in which they were introduced or practised. Examples of items in this cluster include:

- Interpreting a graph modelling the growth of the combined weight of fish in a waterway over a nine-year period; ascertaining how many years a fisherman should wait to maximise the number of fish caught annually; and providing an argument to support the answer; and
- Interpreting data on the increase in a national defence budget from the competing perspectives of a pacifist and a military person.

Examples of items in Appendix A that are in the Reflection cluster include Exchange Rate Q3, Internet Relay Chat Q2, and Earthquake Q1.

Mathematical Situations and Contexts

Engagement with mathematics – the ability to use and do mathematics in a variety of situations – is viewed as being an important part of mathematics. The type of mathematics or method employed often depends on the situation in which the problem is presented. The PISA mathematics framework identifies four situation-types which it is believed students encounter in their everyday lives – personal, educational/occupational, public, and scientific. These situations are designed to be 'authentic' to the extent that the use of mathematics is genuinely directed at solving the problem at hand, rather than the problem merely being a vehicle for the purpose of practising mathematics.

While situation indicates the part of the student's world to which the problem belongs, context reflects the specific setting within that situation. In other words, the context provides the necessary information required to solve the problem. For example, a context might concern the different interest rates offered by a bank.

The framework distinguishes between contexts which are intra-mathematical and those which are extra-mathematical. Questions which do not move beyond the mathematical realm are said to be intra-mathematical, and therefore are considered mathematical situations. Conversely, extra-mathematical problems go beyond an immediate mathematical context. They are concerned with real-world objects and everyday experiences and may not be explicitly mathematical. Students should recognise that they are being presented with a mathematics question and interpret it appropriately in order to extract the relevant information.

Characteristics of the PISA 2003 Mathematics Item Set

In PISA 2003, there were 54 mathematics units. Each unit consisted of a problem context – usually a brief written description of the problem, and associated graphics – and one or more items. There were 85 mathematics items or tasks. Table 1.2 shows how these were distributed across different dimensions of the framework. Items are almost evenly distributed across the four overarching ideas. In the case of the competency clusters, however, there are proportionately fewer items in the Reflection cluster (22.4%) and proportionately more in the Connections cluster (47.1%), with 30.6% in the Reproduction cluster.

Table 1.2. Distribution of PISA 2003 Mathematics Items by Dimensions of the Mathematics Framework

Dimension	*Number of Items*	*Percent of Items*	*Dimension*	*Number of Items*	*Percent of Items*
Overarching Ideas			***Competency Cluster***		
Space & Shape	20	23.5	Reproduction	26	30.6
Change & Relationships	22	25.9	Connections	40	47.1
Quantity	23	27.1	Reflection	19	22.4
Uncertainty	20	23.5	Total	85	100.0
Total	85	100.0			
			Item Type		
Situation			Simple Multiple Choice	17	20.0
Personal	18	21.2	Complex Multiple Choice	11	12.9
Educational/Occupational	21	24.7	Short Response Items	23	27.1
Public	29	34.1	Closed Constructed Resp	13	15.3
Scientific	17	20.0	Open Constructed Resp	21	24.7
Total	85	100.0	Total	85	100.0

Multiple-choice questions are viewed as being most suitable for assessing processes associated with the Reproduction and Connections competency clusters (though not all these items are multiple-choice), while open-constructed response items, which allow for a more extended answer involving higher-order processes, are more likely to be associated with the Reflection cluster. In PISA 2003 mathematics, one-fifth of the items were simple multiple-choice, 12.9% complex multiple-choice items (such as a series of yes/no questions within an item), 27.1% short response, 15.3% closed constructed response, and 24.7% open constructed response (Table 1.2). Partial credit was available for partially correct responses for 10 of the open constructed response items. Double-digit coding, which entails scoring items for both correctness and error patterns, was used for nine items. Examples of item types and scoring keys used in the PISA 2003 assessment may be found in Appendix A.

In addition to categorising the mathematics items by three dimensions (overarching idea, competency cluster, situation), the PISA framework also categorised items according to more traditional mathematics strands. Table 1.3 provides a cross-tabulation of overarching ideas and mathematical strands.

Table 1.3. Cross-tabulation of PISA 2003 Mathematics Items by Overarching Idea and Strand Area

	PISA Overarching Idea									
	Space and Shape		*Change and Relationships*		*Quantity*		*Uncertainty*		*Total*	
PISA Strand Area	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Algebra	0	0.0	3	13.6	0	0.0	0	0.0	3	3.5
Discrete Maths	0	0.0	0	0.0	4	17.4	1	5.0	5	5.9
Functions	0	0.0	9	40.9	0	0.0	0	0.0	9	10.6
Geometry	18	90.0	0	0.0	0	0.0	0	0.0	18	21.2
Number	2	10.0	5	22.7	19	82.6	1	5.0	27	31.8
Probability	0	0.0	0	0.0	0	0.0	5	25.0	5	5.9
Statistics	0	0.0	5	22.7	0	0.0	13	65.0	18	21.2
Total	20	100.0	22	100.0	23	100.0	20	100.0	85	100.0

Ninety percent of Space & Shape items are categorised as belonging to the strand area of geometry, while the same percentage of Uncertainty items are in the area of probability or statistics. Similarly, 82.6% of Quantity items are in the area of number. However, items in the Change & Relationships subdomain are distributed over algebra (13.6%), functions (40.9%), number (22.7%), and statistics (22.7%). In Chapter 6, consideration is given to the extent to which the overarching ideas and strand (content) areas relate to the Junior Cycle mathematics syllabus topic areas.

Reporting Outcomes for Mathematics in PISA 2003

The mathematics framework was more extensive than in PISA 2003 than in PISA 2000, in which mathematics was a minor assessment domain. In 2000, the mathematics assessment comprised 32 items distributed over two over-arching ideas (Growth & Change[3], Space & Shape). Unlike PISA 2000, PISA 2003 has a sufficiently large item pool to allow reporting by mathematics subdomain, with separate scales for Space & Shape, Change & Relationships, Quantity, and Uncertainty, as well as an overall (combined) mathematics scale.

In line with its status as a major domain in PISA 2003, a variety of scores are reported for mathematics. First, mean country scores and performance at key markers are reported for each overarching idea (the subscales), and for combined mathematics (the combined scale). Second, proficiency levels are described for the four subscales and for the overall scale. Third, the mean scores on mathematics for students in 2000 and 2003 are compared for two sub-scales: Space & Shape and Change & Relationships.

FRAMEWORK FOR READING LITERACY

The reading literacy assessment in PISA 2003 did not measure whether 15-year old students were technically able to read (i.e. whether or not they could recognise words in written text). Rather, assuming that they could, it aimed to assess the ability of students to understand and reflect on a wide range of written materials in a variety situations in which students were likely to encounter such materials at and beyond school. Reading literacy is defined as

> …understanding, using and reflecting on written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society. (OECD, 2003, p. 15)

In addition to basic understanding of texts, this definition draws attention to higher-order reading comprehension skills including using and reflecting on texts. Reference to participation in society emphasises the role of reading literacy in economic, political, cultural and social life, and reading in later life (life-long learning).

In operationalising this definition, reading literacy is defined in terms of three dimensions: the content and structure of texts; the processes that need to be performed; and the situations in which knowledge and skills are drawn on or applied.

[3] Growth & Change in PISA 2000 is equivalent to Change & Relationships in PISA 2003.

Text Content/Structure

PISA assesses students' understanding of two text types – continuous and non-continuous. Continuous texts consist of sentences arranged in paragraphs, which, in turn, may be arranged into longer texts such as sections, chapters or books. Non-continuous texts are frequently organised in two-dimensional matrix format, based on combinations of lists, and include charts and timetables.

Continuous and non-continuous texts are further subdivided into 11 text types, of which seven were assessed in PISA 2003. The first three of these are examples of continuous texts, while the remainder are examples of non-continuous texts. The text types are:

(i) Description (of persons, places or objects);
(ii) Narration (stories, reports, news articles);
(iii) Exposition (essays, definitions, explications, summaries);
(iv) Charts and graphs;
(v) Forms;
(vi) Maps;
(vii) Tables.

Reading Processes

PISA identifies three reading processes:

(i) Retrieving information – locating one or more pieces of information in independent parts of a text;
(ii) Developing an interpretation – constructing meaning and drawing inferences using information from one or more parts of the text; and
(iii) Reflecting on the content and form (structure) of texts – relating a text to one's experience, knowledge and ideas, to provide a critical evaluation.

These reading processes are not viewed as constituting a hierarchy. Rather, it is recognised that each may require the same underlying skills, and that readers at each developmental level may draw on at least some aspects of the three processes as they comprehend a text.

As no new reading literacy items were released into the public domain following PISA 2003, readers may wish to look at items that were released following the 2000 assessment (see Cosgrove, Sofroniou, Kelly, & Shiel, 2003; OECD, 2003; Shiel et al., 2001)[4]. In Cosgrove et al. (2003), a narrative text, the Gift, includes items categorised as assessing each of the PISA reading processes.

Reading Situations

Situation refers to the uses of, and purposes for which, texts were constructed. The situations in which reading takes place, defined as how the author intended the text to be used, include:

(i) reading for private use (personal) – reading to satisfy one's own interests, whether practical, emotional or intellectual;
(ii) reading for public use – reading for participation in the activities of the wider society;

[4] Reading literacy texts and items released into the public domain after the PISA 2000 assessment may also be viewed at http://www.erc.ie/pisa.

(iii) reading for work (occupational) – reading of the type encountered in occupational settings; and

(iv) reading for education – reading to acquire information as part of a broader learning task.

Characteristics of the Reading Literacy Item Set

In PISA 2003, the three framework dimensions (content/structure, process, and context) were represented in a series of texts (eight in all), and a number of tasks (28 items), drawn from the much larger pool of texts and items used in PISA 2000 (in which reading literacy was the major domain). As in 2000, almost two-thirds (64.3%) of the items in 2003 are based on continuous texts, and one-third (35.7%) on non-continuous texts (Table 1.4). Similarly, the proportions of items categorised as Retrieve (25.0%), Interpret (50.0%), and Reflect/Evaluate (25.0%) in 2003 are almost identical to 2000. Unlike 2000, none of the 2003 items is drawn from the text types classified as advertisements, argument/persuasive, or schematics. In 2003, each of the reading contexts is represented by between 20% and 30% of the items.

The distributions of item types across the 2000 and 2003 reading literacy assessments are also broadly similar. In 2003, 32.1% of items were of the simple multiple-choice variety, 3.6% complex multiple-choice, 14.3% short response, 14.3% closed constructed response, and 35.7% open-constructed response (Table 1.4). Open-constructed response items are used more often to assess processes in the Reflect/Evaluate category than in the Interpret or Retrieve categories, where item types are more evenly distributed.

Table 1.4. Distribution of PISA 2003 Reading Literacy Items by Dimensions of the Reading Literacy Framework

Dimension	*Number of Items*	*Percent of Items*	*Dimension*	*Number of Items*	*Percent of Items*
Text Structure			***Reading Situations***		
Continuous	18	64.3	Personal	6	21.4
Non-continuous	10	35.7	Public	7	25.0
Total	28	100.0	Occupational	7	25.0
			Educational	8	28.6
Text Type*			Total	28	100.0
Narrative (C)	3	10.7			
Expository (C)	12	42.9	***Item Type***		
Descriptive (C)	3	10.7	Simple M. Choice	9	32.1
Charts and graphs (NC)	2	7.1	Complex M. Choice	1	3.6
Tables (NC)	4	14.3	Short Response	4	14.3
Maps (NC)	1	3.6	Closed Con. Response	4	14.3
Forms (NC)	3	10.7	Open Con. Response	10	35.7
Total	28	100.0	Total	28	100.0
Reading Process					
Interpreting	14	50.0			
Reflecting/Evaluating	7	25.0			
Retrieving information	7	25.0			
Total	28	100.0			

(C) denotes continuous text; (NC) denotes non-continuous text.
*Some text types, i.e. advertisements, arguments, injunctives, and schematics were assessed in 2000 but not in 2003.

Reporting Outcomes for Reading Literacy in PISA 2003

In PISA 2000, overall performance on reading literacy was reported in terms of a combined (overall) reading literacy scale and five subscales – two based on text type (continuous/non-continuous), and three based on text processes (retrieve, interpret, reflect/evaluate). In PISA 2003, performance on reading literacy is reported with reference to the combined reading literacy scale only. Country mean scores are given, and the proportions of students who achieve each proficiency level on the same scale are also given. Performance in 2003 is also compared with performance in 2000. Again, mean scores and percentages of students scoring at each proficiency level on the combined reading literacy scale are compared.

FRAMEWORK FOR SCIENTIFIC LITERACY (SCIENCE)

PISA is concerned with the capacity of students to draw appropriate and guarded conclusions from evidence and information given to them, to criticise the claims made by others on the basis of the evidence, and to distinguish opinion from evidence-based statements. The ability to apply such processes in the context of science gives rise to the concept of ‘scientific literacy’. According to PISA, scientific literacy (science) is

> …the capacity to use scientific knowledge, to identify questions and to draw evidence-based conclusions, in order to understand and help make decisions about the natural world and the changes made to it through human activity. (OECD, 2003, p. 133)

The use of the term ‘scientific literacy’ implies that both scientific knowledge (knowledge about science) and knowledge of processes are important. Indeed, the ability to apply scientific processes is viewed as requiring scientific knowledge, in the same way as the ability to engage in mathematical processes (mathematisation) is thought to require knowledge of mathematical terminology, facts, and procedures.

The PISA science framework includes three dimensions: scientific knowledge or concepts; scientific processes; and scientific areas of application (situations).

Scientific Knowledge or Concepts

The concepts included in PISA are selected from the four major fields of physics, chemistry, biological science, and Earth and space science, based on three criteria: the relevance of the knowledge/concept to everyday life; the likely relevance of the knowledge/concept in the next decade and beyond; and whether the knowledge can be combined with selected scientific processes. The scientific knowledge and concepts assessed in PISA 2003 include: structure and property of matter; atmospheric change; chemical and physical changes; energy transformations; forces and movement; form and function; physiological change; genetic control; ecosystems; Earth and its place in the universe; and geological change.

Scientific Processes

Processes such as explaining, understanding, or interpreting become scientific when subject matter is drawn from the world of science, and the outcome of applying processes is to further scientific understanding. With its ‘science for all’ focus, PISA science emphasises the skills and processes needed to collect and interpret evidence from the world around us, and to draw conclusions from it. Thus, within PISA, ‘priority is given to processes *about* science as compared to processes *within* science’ (OECD, 2003, p. 137).

The PISA framework distinguishes between three major scientific processes:

(i) *Describing, explaining, and predicting scientific phenomena*. This involves demonstrating understanding by applying the appropriate scientific knowledge in a particular situation, describing or explaining scientific phenomena and predicting changes or outcomes;

(ii) *Understanding scientific investigation.* This involves having a grasp of the type of questions that can be answered by scientific investigation, and the ability to communicate them. The process also involves identifying the evidence required, the information that should be collected and the variables that are involved; and

(iii) *Interpreting scientific evidence and conclusions.* This refers to understanding the findings from scientific investigations. Students should be able to identify relevant evidence or data and derive conclusions, selecting from a range of explanations and developing arguments for and against the various conclusions that are possible. This process can also entail reflecting on the impact or implication of the conclusion reached.

Scientific knowledge is required for all three processes. It is viewed as being essential for the first, and necessary, but not sufficient, for the second and third, since knowledge about collecting and using scientific evidence and data is necessary for understanding scientific investigation and interpreting scientific evidence and drawing conclusions, but other interpretative processes are also involved.

Sample items accompanying the assessment framework for PISA 2003 (OECD, 2003) show how both knowledge of science and scientific processes are assessed in the same unit. In a unit called ‘Corn’, for example, students are asked to read a passage about the potential use of corn as a fuel, and are asked to complete an equation for photosynthesis. Students are also asked to evaluate the adequacy of data used by a farmer to argue that carbon dioxide is not the main cause of the greenhouse effect. While the former item emphasises scientific knowledge, the latter focuses on interpreting scientific evidence and drawing conclusions.

The PISA 2003 science categories differ somewhat from those underpinning PISA 2000, though both sets seem to assess the same underlying processes.[5]

Scientific Areas of Application (Situations)

As with mathematics and reading literacy, three broad situations in which individuals apply scientific processes were identified:

(i) Science in life and health, including health, disease and nutrition, maintenance and sustainable use of species, and interdependence of physical/biological systems;

(ii) Science in Earth and environment, including pollution, production and loss of soil, and weather and climate; and

[5] The first and second PISA 2000 science processes, ‘recognising scientifically investigable questions’ and ‘identifying evidence needed in a scientific investigation’ are similar to the PISA 2003 science process of ‘understanding scientific investigation’. The third PISA 2000 science process, ‘demonstrating understanding of scientific concepts’ is similar to the PISA 2003 science process of ‘describing, explaining and predicting scientific phenomena’, although some of the items in this category might be classified as ‘interpreting scientific evidence and conclusions’. The fourth and fifth PISA 2000 science process, ‘drawing or evaluating conclusions’ and ‘communicating valid conclusions’ are subsumed by the PISA 2003 process, ‘interpreting scientific evidence and conclusions’.

(iii) Science in technology, including biotechnology, use of materials and waste disposal, use of energy, and transportation.

As no new science items were released following PISA 2003, readers may wish to look at items that were released following the 2000 assessment (see OECD, 2003; Shiel et al., 2001)[6].

Characteristics of the Science Item Set

The elements of the framework were brought together in 2003 in a set of 13 science units containing 35 test items (Table 1.5). The distribution of items is broadly similar to 2000, when there were also 35 items (though the actual number of items common to both assessment cycles is 25). One significant change is the collapsing of five scientific processes in 2000 into three in 2003. Nevertheless, lower level processes (describing, explaining and predicting scientific phenomena) are again well represented, with 48.6% of items tapping such processes. The higher-level process of understanding scientific investigation is represented by 31.4% of items, while understanding scientific evidence and conclusions is represented by 20.0%. As in 2003, Earth and biological sciences are more strongly represented in the item pool than the physical sciences.

Table 1.5. Distribution of PISA 2003 Science Items by Dimensions of the Science Framework

Dimension	*Number of Items*	*Percent of Items*
Scientific Knowledge/Concept		
The Earth and its place in the universe	7	20.0
Structure and property of matter	6	17.1
Energy transformations	4	11.4
Physiological change	4	11.4
Atmospheric change	3	8.6
Form and function	3	8.6
Ecosystems	3	8.6
Genetic control	2	5.7
Chemical and physical changes	1	2.9
Forces and movement	1	2.9
Geological change	1	2.9
Total	35	100.0
Scientific Process		
Describing, explaining and predicting scientific phenomena	17	48.6
Understanding scientific investigation	11	31.4
Interpreting scientific evidence and conclusions	7	20.0
Total	35	100.0
Scientific Areas of Application (Situation)		
Science in life and health	12	34.3
Science in Earth and in the environment	12	34.3
Science in technology	11	31.4
Total	35	100.0
Item Type		
Simple Multiple-Choice	13	37.1
Complex Multiple-Choice	7	20.0
Short Response	1	2.9
Open Constructed Response	14	40.0
Total	35	100.0

[6] Science texts and items released after the PISA 2000 assessment may be viewed at http://www.erc.ie/pisa

The item types in PISA 2003 were simple multiple-choice (37.1%), complex multiple-choice (20.0%), short response (2.9%) and open constructed response (40.0%). There were no closed constructed response items.

The scientific situations of science in life and health, science in Earth and in the environment, and science in technology are evenly represented. As in other domains, these situations also reflect the concern with students' future lives that permeates PISA.

Reporting Outcomes for Science in PISA 2003

As in PISA 2000, performance in science in 2003 is reported with reference to an overall scale only. There are no subscales, and proficiency levels are not used, although the scientific knowledge and skills expected of students at selected markers on the overall scale are given. Mean scores of countries on the overall scales in 2000 and 2003 are compared. When science becomes the major assessment domain in 2006, it will be possible to obtain more detailed information about the performance of students on the components (subdomains) of scientific knowledge, and on different scientific processes. It is also planned to assess students' attitudes towards science.

FRAMEWORK FOR PROBLEM SOLVING

As the frameworks for mathematics and science suggest, the ability to engage in domain-specific problem solving is viewed as an important learning outcome in PISA. Reflecting its interest in tapping into competencies believed to impact on performance across a range of subject domains, PISA also developed an assessment of cross-curricular problem solving for 2003. The assessment, which is not expected to be included in future PISA cycles, was given the status of a minor domain. A key issue in interpreting the outcomes of the assessment of problem solving is the extent to which it is independent of subject-specific domains, including mathematics. This question is taken up in Chapter 7.

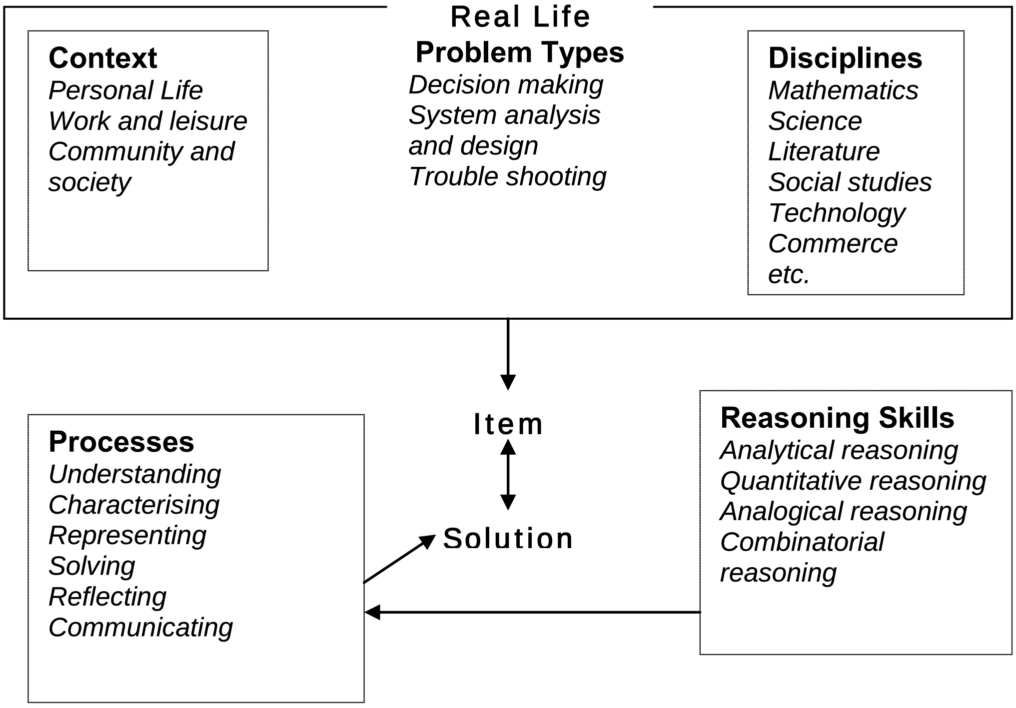
The literature on cognitive psychology includes small-scale efforts to examine various aspects of problem solving including knowledge of inductive reasoning, analogical reasoning, and analytical problem solving (e.g., Csapó, 1997; Vosniadou & Orthony, 1989), while, as part of PISA 2000, the problem-solving skills of 650 15-year-olds in Germany were assessed (Klieme, 2000). An attempt was made to extend the prototypes developed in earlier research and feasibility studies to a workable model for a large-scale assessment as part of PISA 2003.

Although there is no standard definition of cross-curricular problem solving, it is generally acknowledged that the major focus should be on the solver's ability to attempt a problem, search for a solution, and finally communicate the results of these activities. With this in mind, problem solving in PISA 2003 is defined as follows:

> …an individual's capacity to use cognitive processes to confront and resolve real, cross-disciplinary situations where the solution path is not immediately obvious and where the literacy domains or curricular areas that might be applicable are not within a single domain of mathematics, science or reading. (OECD, 2003, p. 156)

Cognitive processes here refer to understanding, characterising, representing, solving, reflecting, and communicating. The term 'cross-disciplinary' indicates that PISA problem solving extends across subject domains, including but not confined to those already represented in PISA. The term 'real' highlights the importance of situating problem-solving tasks in contexts arising from real-life situations.

Figure 1.3. Key Components of the PISA 2003 Problem-Solving Framework



Source: OECD (2003), Figure 4.1, p. 159.

The problem-solving framework is organised according to the type of problem encountered, the processes involved in solving a problem, and the situations involved (Figure 1.3).

Problem Types

PISA problem solving involves 10 units comprising 19 items, which are categorised according to three problem types: decision-making; system analysis and design; and trouble shooting. It is argued that 'the three… problem types provide the generic structures within which problem-solving processes can be assessed' (OECD, 2003, p. 160).

Decision-Making

Decision-making problem-solving tasks require students to understand the situation presented in a problem which involves a number of constraints, select from a number of alternatives presented, and make a decision in keeping with the constraints. The complexity of problems increases with the amount of information given, and/or the number of criteria upon which to base a decision. Once made, the results of the decision must be examined and then communicated to an audience. To solve such problems, students must combine information from a number of diverse sources (combinatorial reasoning) and select the best solution. As part of such problems, students may be presented with graphs or tables and asked to interpret them, or they may be asked to create relevant representations. One example of a decision-making task involves reading the directions on four packets of medication, and making decisions regarding the appropriate one for a range of symptoms and age groups (see OECD, 2003). Another, Cinema Outing, which requires students to make a choice about which films to go to in the cinema, given a number of criteria and constraints, may be found in Appendix A (see also OECD, 2004c).

System Analysis and Design

This problem type requires students to analyse a complex situation in order to understand it and/or to design a system to achieve a certain outcome, given information about features of the problem context. This type of problem solving differs from decision-making in two ways. Firstly, in most cases, students must design a solution rather than select from a set of alternatives. Secondly, the situation usually involves a complex set of interrelated variables. Variables in decision-making problems do not usually interact. Here, however, the variables have dynamic relationships and the difficulty of the item is determined not only by the number of variables but also by the complexity of their interrelationships. As with decision-making, the solution must be evaluated, justified, and then communicated to an audience. An example of a system analysis and design problem is to identify variables that are relevant in establishing a system for tracking compact disc sales in a music store, taking into account year, publisher, artist, and serial number (see OECD, 2003). Another example, Library Systems, which asks students to examine a system for borrowing and tracking library books, may be found in Appendix A in this report (see also OECD, 2004c).

Troubleshooting

This problem type involves understanding the way a system functions in order to identify a fault or underperforming feature. In this instance, there are no alternative solutions presented and no systems to design. Rather, students are required to understand an existing system, identify the essential features, locate the problem, and suggest solutions. Processes involved in troubleshooting include understanding the logic of a causal mechanism, identifying the critical features for the diagnosis of a specific problem, creating or applying relevant representations, diagnosing the problem, proposing a solution, and, where relevant, executing a solution. An example is identifying the fault in a bicycle pump, given information on its typical operation (see OECD, 2003). Another example, Irrigation, may be found in Appendix A (see also OECD, 2004c).

Problem-Solving Processes

It is difficult to provide a definitive list of the processes involved in problem solving as individuals can use a variety of novel methods. The skills underlying the six processes described here are considered to be those most closely related to the three problem types encountered in PISA problem-solving items. Moreover, there is strong support for them in the literature on problem solving (e.g., Polya, 1945). There is no implied hierarchy, and not all six processes are necessarily involved in solving every problem.

(i) *Understanding the problem.* Understand and draw inferences from the information presented. The process includes the ability to incorporate material from different sources and to use this along with any relevant background knowledge to comprehend the problem;

(ii) *Characterising the problem.* Identify the variables presented and establish how they are inter-related. The process includes determining which variables are relevant and constructing hypotheses based on the contextual information;

(iii) *Representing the problem.* Represent the problem in various formats (e.g., tabular, graphical, verbal). The process includes the ability to move from one representational format to another;

(iv) *Solving the problem.* Make a decision or analyse/design a system or diagnose and provide a solution, depending on which of the three problem types is encountered;

(v) *Reflecting on the solution.* Examine and evaluate the solution. This involves the ability to adopt a variety of perspectives and, if necessary, to make adjustments and then justify the solution chosen; and

(vi) *Communicating the problem solution.* Select appropriate media and representations to best communicate the solution chosen.

Problem Situations

The problems in PISA are embedded in real-life situations. The cross-disciplinary nature of problems means that, in many cases, they approach the notion of 'life skills'. Problems are embedded in real-life settings associated with personal life, work or leisure, or community and society.

Design of the Problem-Solving Item Set

Problems in PISA are categorised by problem-solving type, with 36.8% of items categorised as decision-making, 36.8% as systems analysis and design, and 26.3% as troubleshooting (Table 1.6). Thirty-seven percent of items have a multiple-choice format, 47.4% an open constructed response format, and 15.8% a closed constructed response format. The problem-solving items are not classified further in terms of process or situation.

Table 1.6. Distribution of PISA 2003 Problem-Solving Items by Dimensions of the Problem-Solving Framework

Dimension	*Number of Items*	*Percent of Items*
Problem Type		
Decision-Making	7	36.8
System Analysis and Design	7	36.8
Trouble Shooting	5	26.3
Total	19	100.0
Item Type		
Simple Multiple-Choice	3	15.8
Complex Multiple-Choice	4	21.1
Open Constructed Response	9	47.4
Closed Constructed Response	3	15.8
Total	19	100.0

Reporting Outcomes for Problem Solving in PISA 2003

Given the relatively small number of items (19 in all), performance on problem solving in PISA 2003 is reported in terms of a single scale. Descriptions of the processes that students scoring at different proficiency levels along the scale are likely to be able to perform are also given.

THE PISA CONTEXT QUESTIONNAIRES

Across all participating countries in PISA 2003, information about the backgrounds of students and their schools was obtained from the students themselves and from their principal teachers through the administration of Student and School Questionnaires. Data from the questionnaires were subsequently used to interpret students' performance on the assessments of mathematics, reading literacy, science, and problem solving. The Student and School Questionnaires are described in this section.

The PISA questionnaires were set in the broader context of the OECD INES (Indicators of National Educational Systems) project and sought to obtain information on the social, cultural, economic, and educational factors that are associated with student achievement. The specific content of the questionnaires was informed by the policy issues identified by participating countries. In PISA 2003, these included students' attitudes towards mathematics and their educational pathways.

The PISA questionnaire framework provides a conceptual overview of variables associated with achievement, which include the antecedents of learning, the contexts in which learning occurs, and the content of learning. These are considered at four levels: the system, the school, the class and the student (Figure 1.4). It is implied in Figure 1.4 that the flow of information in the model is bi-directional, with different levels influencing each one another. Areas of the framework that were covered in detail in the questionnaires include school conditions, student background, and student classroom behaviours (with an emphasis on behaviour in mathematics classes). Information was also gathered on students' engagement in mathematics. In Ireland, information regarding instructional practices and classroom learning processes was obtained in a Teacher Questionnaire; results will be described in a future publication.

Figure 1.4. PISA Questionnaire Framework

	Antecedents	*Contexts*	*Contents*
System	Country features (e.g., GDP*, wealth distribution)**	Instructional settings and policies (e.g., school management, teacher qualifications and training)	Intended schooling outcomes (statements of goals for teaching and learning)
School	Community and school features (e.g., school environments and practices; location, size, structure, management)	School conditions and processes (e.g., school equipment, teacher qualifications, school climate variables)	Implemented curriculum
Class	Teacher background	Class conditions, processes (climate, use of homework, class size, instruction time)	Implemented curriculum
Student	Student background (e.g., socioeconomic status, parental occupation, wealth, cultural capital, age, gender, school attendance)	Student classroom behaviour (e.g., attendance, engagement in work, confidence in own mathematics skills)	Attained schooling outcomes (mathematics, reading literacy, science, attitudes to mathematics)

*GDP = Gross Domestic Product. This is an indicator of a country's total production and expenditure, and reflects the activities of foreign-owned corporations as well as domestically-owned corporations operating in a country.

**System-level information on the antecedents of learning was obtained from sources other than PISA.
Note. Shaded cells indicate aspects of education systems that are addressed in the PISA School and Student Questionnaires. Those in white are not examined directly by PISA.

It should be acknowledged that, while each cell in the framework has a conceptual basis, it is unclear how the cells relate to one another. This is partly a result of the complexity of variables and relationships that potentially influence student learning outcomes. It is also partly due to PISA's cross-sectional design, which does not permit causal inferences. Further, since PISA students are randomly sampled from within schools rather than within classrooms, and therefore come from different learning environments with different teachers, and different levels of instruction, PISA does not offer an international-level teacher questionnaire. Nevertheless, data on class-level variables such as teacher expectations were collected from principal teachers and/or students.

While all countries participating in PISA 2003 were required to administer a core set of questions to both students and principal teachers, countries had the option of including additional questions that were of national interest only. The subsections that follow describe the core and optional content of the Student and School Questionnaires for Ireland.

Student Questionnaire

All students who responded to the PISA cognitive assessments were asked to complete the PISA 2003 Student Questionnaire. Questions that all countries were required to ask of students covered the following topics:

- Background variables, including gender, family structure, socioeconomic status, and parental education;
- Home educational climate, including level of home educational resources and number of books in the home;
- Student as learner, including attitudes to school, sense of belonging at school, and time spent on homework and study;
- School characteristics, including student-teacher relations, disciplinary climate, and teacher support for learning;
- Students' engagement with mathematics, including interest in and enjoyment of mathematics, anxiety about mathematics, learning strategies for mathematics, and mathematics self-efficacy; and
- Educational career, including grade repetition.

A number of countries, including Ireland, also administered an optional international question set on frequency of use of, and attitudes towards, information and communications technology (ICT). Results will be reported in a later publication.

A number of questions of national interest that were included in the Student Questionnaire were identified by the PISA National Advisory Committee because of their relevance for research and policy. They include:

- Whether the student was a member of the Traveller or settled community;
- The year in which the student attempted or intended to attempt the Junior Certificate Examination for the first time;
- The level at which the student studied English, mathematics, and science (if applicable) for the Junior Certificate Examination; and
- Student’s reading interests and frequency of reading a range of reading materials.

Because of the large number of items on the Student Questionnaire, only selected variables are looked at in this report. Some of the variables are composites, based on student responses to a number of related items. Original and composite student-level variables used in this report are described in Appendix B, Section B.1.

School Questionnaire

Principal teachers were requested to complete the PISA 2003 School Questionnaire. This was required in all participating countries and was designed to elicit information on:

- School structure, including enrolment size, admissions policy, length of school year and school day, decision-making procedures, sources of funding, student assessment procedures;
- School climate/policy, including school autonomy, student behaviour, teacher expectations, and teacher and student morale;
- School resources, including student-teacher ratio, class size, computer-student ratio, and variables perceived to hinder the school's capacity to provide instruction; and
- School's strategies to promote engagement with mathematics in the school, and provision for lower-achieving students in mathematics.

Questions identified as important by the PISA National Advisory Committee and administered to school principals in Ireland related to:

- School policy for students with first languages other than English/Gaeilge;
- Provision of extra tuition in the language of instruction to students whose first language is not English or Gaeilge (Irish);
- Availability of the Transition Year programme in the school;
- Percentage of students transferring to another school before the end of Junior Cycle;
- Percentage of students leaving school before the end of Junior Cycle; and
- Percentage of Third year students receiving learning support in English and/or mathematics.

In addition to this information, school-level data from the databases of the Department of Education and Science and the State Examinations Commission were obtained. These included the status of the school with respect to disadvantage (designated disadvantaged status), and the proportion of students in receipt of a fee waiver for the Junior Certificate Examination in 2002 and 2003.

Again, it was not possible to present the responses of principal teachers to all the items on the School Questionnaire in the current report. Some of these will be addressed in future reports. A more complete description of the school-level variables used in the current report, including composite variables, variables derived from student responses, and variables based on information obtained from other sources, may be found in Appendix B, Section B.1.

IMPLEMENTATION OF PISA 2003 IN IRELAND

In this section, the focus is on the implementation of PISA in Ireland, including the field trial undertaken in 2002 and the main study in 2003. The study was jointly implemented by the Department of Education and Science and the Educational Research Centre (ERC). All phases of implementation (except for the translation of assessment materials into Gaeilge) were adjudicated with reference to internationally agreed standards by the PISA International Consortium, led by the Australian Council for Educational Research (see *Data Adjudication in PISA 2003*, PISA Consortium, 2003; and *Standards for PISA 2003*, PISA Consortium, 2001).

Development of Test Materials and Questionnaire Items

Prior to the field trial, participating countries were invited to develop texts (units) and items (questions) on mathematics, science, and cross-curricular problem solving. Items for reading literacy and some items for mathematics and science were drawn from the PISA 2000 item pool. New items were submitted to the international consortium responsible for developing the PISA assessment materials, and subjected to an initial screening for compatibility with the PISA frameworks. Subject matter specialists in participating countries (i.e., persons with extensive experience in syllabus development/assessment in their subject areas) were then invited to comment on the appropriateness of these materials for assessing 15-year-olds in their countries. Following this, the international consortium prepared item pools for use in the field trial. The items in the PISA context questionnaires were developed using a similar process of consultation and review.

Field Trial

To evaluate the appropriateness of the test and questionnaire items, and to examine the effectiveness of field operation procedures, a field trial was conducted in 2002 in all countries intending to participate in PISA 2003. Its purpose was to familiarise countries with the procedures for administering PISA 2003, and to pilot new test items prepared for the 2003 assessment. In Ireland, a sample of 742 students in 30 randomly-selected schools in Dublin, Cork and Galway participated in the trial. Students' responses to the test questions were scored at the ERC by trained markers (third-level students with specialisations in the PISA domains and/or a background in psychology) using scoring rubrics provided by the international consortium.

Feedback from test administrators (retired school inspectors) indicated that implementation of the field trial went as planned in almost all schools. The international consortium selected items for PISA 2003 based on the outcomes of the field trials. In addition to the relative difficulty of items and other standard psychometric properties, differential item functioning (the extent to which performance on particular items departed from expected patterns for certain subgroups, including males and females) was taken into account in the item selection process.

Main Study

The PISA main study was conducted in Ireland in March 2003. In this section, the target population and exclusions, the sample design, and the administration are described.

Sampling Schools and Students

The target population (i.e., the population from which individuals are potentially eligible for selection) comprised all 15-year old students (those born between January 1 and December 31, 1987) who were in full-time education in recognised schools (school-going total = 59,801). This covered approximately 97.2% of 15-year-olds in Ireland (2.8% were not at school).[7] For logistic reasons, one type of school-level exclusion was made to the national target population in advance of sampling. Five schools on islands, with a total enrolment of 91, were excluded. The national desired population then consisted of 99.85% of the national

[7] This estimate is based on dividing the total number of individuals born in 1987 who were enrolled in recognised schools (both primary and second-level) by the estimated total number of individuals born in 1987. Data for these estimates were taken from Tables 1.2 and 1.4 of the 1999/2000 *Annual Statistical Report* of the Department of Education and Science (2001).

enrolled population (N = 59,710). The defined target population (i.e., the target population after school-level exclusions made for logistic reasons) consisted of students in 'recognised' schools (i.e. mainstream second-level schools in which teaching staff salaries are paid by the Department of Education and Science). The following school-level exclusions were made: non-aided schools (5 schools with an estimated total enrolment of 30 students aged 15), primary schools in which 15-year-olds were enrolled (98 schools with an estimated total enrolment of 119), and special needs schools in which 15-year-olds were enrolled (50 schools with an estimated total enrolment of 785). The estimated total enrolment of 15-year-olds in these schools (934) accounts for 1.6% of the defined target population. Thus, a population coverage index of 98.4% is achieved before within-school exclusions.

A two-stage stratified sample design was used. Schools were selected first, followed by students within schools. In the first stage, schools in the sampling frame were stratified according to the total number of 15-year-olds in the school:

- Stratum 1: Small schools: 1 to 40 15-year-olds (120 schools);
- Stratum 2: Medium schools: 41 to 80 15-year-olds (272 schools); and
- Stratum 3: Large schools: 81 or more 15-year-olds (325 schools).

Within these three strata, schools were ordered by implicit strata consisting of school type (secondary, community/comprehensive, or vocational) and by gender composition (percent of 15-year-olds enrolled that are female; none, 1-33%, 34-66%, 67-99%, all). In PISA 2000, gender composition was indexed by categorising schools as all boys, all girls or mixed. However, it was decided that a measure of the proportion of 15-year old females in a school would yield a more accurate description of gender composition for 2003. To achieve a sample size of 5250 students, as recommended by the international consortium, 155 schools were selected to participate. Within each stratum, schools were selected with probability proportional to size (with school type and gender operating as implicit stratifying variables).

In PISA 2000, schools with an estimated enrolment of 16 or fewer 15-year-olds were excluded. In 2003, to increase population coverage, all schools with 15-year-olds were included, and if a school with a very small number of 15-year-olds was drawn, it was excluded on a *post-hoc* basis after checking on student numbers with the school. One school in Ireland was excluded for this reason.

Four of the schools selected were taking part in another large scale Irish survey. It was decided that these schools would not be approached; in each case, a replacement school (the next school on the sorted school file) was invited to participate. Of the schools selected, 141 agreed to participate, giving a weighted response rate of 90.2%. Four replacement schools also agreed to participate, bringing the weighted school-level response rate after replacement to 92.8%.

In the second stage of sampling, the required number of 15-year old students within each participating school were selected at random (i.e., 35 age-eligible students were sampled with equal probability where enrolment exceeded 35; students were sampled with 100% probability from schools with 35 of fewer such students). Among selected students, functionally disabled students, students with general learning disabilities, students with specific learning disabilities, and those with limited proficiency in the test language (English) could be excluded from the assessment by school principals, using PISA guidelines. In total, 2.9% of sampled students were exempt under these guidelines. This, combined with the whole-school exclusions described above, yields a coverage rate of 95.7% of the national desired target population. After refusals, absences and transfer of students to other schools were taken into account, 3,880 students participated in PISA 2003, yielding a weighted within-school response rate of 82.6%. Both school and student response rates were in line with the standards established for PISA 2003 (a response rate of 85% for initially selected schools, and a student-level response rate of 80%).

Since it was estimated that less than 5% of students in Ireland were receiving instruction in mathematics and science in a language other than English, Ireland was not required to implement official PISA translation and verification procedures. Nevertheless, test administration materials, questionnaires, and the tests of mathematics, science and problem-solving literacy were translated into Irish (Gaeilge) to provide students with the option of responding in either English or Irish (in the same way that such an option is given in state examinations). Of the 145 schools that agreed to participate, 10 were Gaelscoileanna or were located in Gaeltacht areas (i.e., were Irish language medium schools). Six of these schools opted to receive the test materials in both languages and 20 students (0.5%) completed the assessment through Irish.

Administration of Assessments

The PISA assessment was administered to selected students in their own schools by retired inspectors of the Department of Education and Science and retired principal teachers. Testing took place within a two-week period in March 2003. The use of a rotated test design meant that each student was asked to attempt just a portion of the full pool of assessment units and items. Of the 13 test booklets used, 7 contained some reading literacy items, 7 contained some science items, and 7 contained some problem-solving items. All of the booklets included some mathematics items (see Appendix A, Table A.1). Testing time was 120 minutes for the cognitive tests. Students were given 45 minutes to complete the Student Questionnaire. Principal teachers were requested to complete the School Questionnaire, seal it, and to give it to the test administrator in their school, who returned it to the Educational Research Centre.

Retired senior inspectors not otherwise involved in PISA were employed by the PISA Consortium to monitor the testing sessions in 14 of the schools. The inspectors reported directly to the consortium on matters such as suitability of the conditions in which the assessment was carried out, the timing of assessment sessions, and whether or not disruptions occurred during the sessions. In almost all cases, administration of PISA 2003 was judged to be in line with internationally agreed standards.

Following the assessments, students' responses were scored at the Educational Research Centre by trained markers, using detailed marking guides provided by the PISA consortium. As was the case in other countries, it was a requirement that a subset of test booklets be marked four times and returned to the PISA Consortium for inter-rater reliability checks using homogeneity analysis. The results of that analysis (not published at the time of writing) suggest that overall reliability was high, and that the reliability of the Irish marking was higher than average.

ANALYSIS OF PISA DATA

Some of the procedures used by the PISA Consortium to scale achievement data are summarised in this section. Procedures underlying the analyses reported in Chapters 4 to 7 are also summarised. More detailed descriptions of the procedures are provided in the relevant chapters.

Scaling of PISA 2003 Achievement Data

Student achievement was scaled using a one-parameter Item Response Theory (IRT) model (specifically, a mixed coefficient multinomial logit model), which provides an efficient way of summarising data when a rotated test design is used (see, Adams & Wu, 2002). The procedure was applied in three steps: national item calibration; international item calibration; and scale score generation. IRT places item difficulty and student ability on the same metric so that student ability at a specific level can be described in terms of task characteristics of items associated with that level. While the difficulty levels of items were known, student ability was imputed or inferred, since each student had taken only a portion of the assessment tasks. This entailed a selection of five likely achievement scores, called plausible values, for each student and each domain. Such values are random numbers that are drawn from the distribution of scale scores that could be reasonably assigned to each individual (see Mislevy, 1991). Plausible values contain random error variance components and are not optimal as scores for individuals. However, analyses that combine all plausible values can be used to describe the performance of groups of students. Five plausible values were assigned to each student for each overall scale (mathematics, reading literacy, science, and problem solving) and for each mathematics subscale.

Plausible values were produced from country-by-country regressions, based on principal components analyses of dummy-coded student questionnaire variables and student gender, student socioeconomic status, and the achievement of the school attended by the student, as represented by its percent correct score. As the explained variance associated with the regression of principal components (derived from student and school contextual variables) on achievement increases, the spread of the achievement distribution from which plausible values are drawn decreases, and thus the measurement error decreases. This procedure is referred to as conditioning.[8] One important procedural difference in scaling student responses in 2003 (compared to 2000) was that estimates of achievement in the four PISA assessment domains were generated for each student participating in the assessment, regardless of whether or not the student had been asked to attempt items in a particular domain.

Estimating Variance Associated with Achievement

The standard errors associated with mean achievement scores presented in this report were computed in a way that took into account the complex, two-stage, stratified sample design. The software used was WesVar 4.2 (Westat, 2000), which incorporates sampling error into estimates of standard errors by a technique known as variance estimation replication, which involves repeatedly calculating estimates for subgroups of the sample and then computing the variance among these replicate estimates. The particular method of variance estimation used was Fay's Balanced Repeated Replication (BRR) method. BRR is commonly used with multistage stratified sample designs, and usually has two units (in this case, schools) in each variance stratum. Using Fay's method, half of the sample is weighted by a K factor (which must be between 0 and 1; for analyses of PISA data, the factor K was set at 0.5), and the other half is weighted by 2-K.

[8] Full details on the development of achievement scales in PISA 2003 will appear in a forthcoming OECD publication, the *PISA 2003 Technical Manual.*

Conducting Multilevel Explanatory Analyses

A limitation of the correlation coefficients and comparisons of mean scores that describe associations between a single explanatory variable and achievement is that they do not take the inter-relatedness of explanatory variables into account, or the possibility that a link between a variable and achievement may occur because both are related to a third variable that was not considered. In Chapter 5, associations between achievement and a range of variables are examined simultaneously using statistical modelling techniques using the R statistical package (Pinheiro & Bates, 2000). Separate models for three PISA domains, mathematics, reading, and science are presented, and the proportions of variance at the school and class/student levels explained by each model are reported.

2

Achievement and Associated Variables: A Review of Earlier International Studies

The purpose of this chapter is to establish a context in which to examine the outcomes of PISA 2003. First, the achievement outcomes of recent international studies, including PISA 2000, are reviewed for mathematics, reading literacy, and science. Second, student- and school-level variables that are associated with achievement are identified, with particular emphasis on findings arising from national analyses of the PISA 2000 data set. Third, analyses of national curricula and links between curricula and performance on international assessments, including PISA 2000, are described.

In general, care should be exercised in interpreting the outcomes of international studies. A country's overall rank order is a rough indication of performance. It does not tell us: (i) whether the performance of a country is significantly different from that of another country or from the international average (taking sampling and measurement error around mean scores into account); (ii) how achievement is distributed within a country (e.g., what proportions of students have high and low achievement levels); or (iii) what variables account for reported patterns of achievement.

The complexity of comparing achievement differences across international studies was revealed in one study in which performance differences among Irish post-primary students in international studies of mathematics and science in 1991 and 1995 were investigated (O'Leary, Kellaghan, Madaus, & Beaton, 2000). A number of features of the assessments that need to be taken into account in comparisons were identified. These include: differences in the target populations (including whether a study uses a grade- or age-based sample); whether or not populations with high exclusion rates and/or low participation in education are included in the data pool; differences in approaches to scaling achievement; actual differences in achievement over time; and differences in the measurement instruments used.

Table 2.1. International Assessments of Achievement in which Ireland Participated (1980-2000)

Year	*Study*	*Areas Assessed*	*Population*
1980-82	Second International Mathematics Study (SIMS)*	Mathematics	1st and 6th years
1989	International Assessment of Educational Progress I	Mathematics, Science	13-year-olds
1991	International Assessment of Educational Progress II	Mathematics, Science	9- and 13-year-olds
1991	IEA Reading Literacy Study	Reading Literacy	9- and 14-year-olds (3rd class, 2nd year)
1994	International Adult Literacy Survey (IALS)	Reading Literacy (including quantitative literacy)	Adults (16- to 64-year-olds)
1995	Third International Mathematics and Science Study (TIMSS)	Mathematics, Science	3rd/4th classes 1st/2nd years
2000	Programme for International Student Assessment (PISA)	Reading, Mathematical, and Scientific literacy	15-year-olds

* Ireland participated in the curriculum analysis component of SIMS (see Oldham, 1989). Achievement data were gathered only in the context of a follow-up study involving students in First year (Carey, 1990) and were not analysed at international level.

Ireland has taken part in several international studies of achievement since the early 1980s, including studies involving primary-level pupils, second-level students, and adults (Table 2.1). In this chapter the focus is on findings of the most recent studies involving school-based populations, while reference is made to earlier studies where relevant.

ACHIEVEMENT OUTCOMES

International Assessments of Mathematics

Between 1980 and 2000, Ireland participated in five international assessments that involved mathematics[9] which used school-based populations: the Second International Mathematics Study (SIMS) in 1980, which involved students in First and Sixth years (second level); the first International Assessment of Educational Progress (IAEP I) in 1988 which involved 13-year-olds; the second International Assessment of Educational Progress (IAEPII) in 1991, which involved 9- and 13-year-olds; the Third International Mathematics and Science Study (TIMSS) in 1995, which involved pupils in Third and Fourth classes in primary schools and in First and Second years in second-level schools; and the Programme for International Student Assessment (PISA) in 2000, which involved 15-year-olds in second-level schools. Of these studies, only SIMS did not yield internationally comparable measures of achievement for Ireland. Another study, the International Adult Literacy Survey (1994), provided a measure of 'quantitative literacy' among adults in the 16-64 years age range.[10] This review focuses on the two most recent studies involving school-based populations – TIMSS (1995) and PISA 2000.

TIMSS Mathematics (1995)

The target populations in the Third International Mathematics and Science Study (TIMSS) were the two adjacent class levels in which the majority of 9- and 13-year olds were enrolled. In Ireland, these were Third and Fourth classes in primary schools and First and Second years (Grades 7 and 8) in post-primary schools. The results for Grade 8 are considered here as this grade level has some overlap with the PISA population (15-year-olds). Forty-one countries participated in TIMSS 1995 assessment at Eighth grade, though just 25 of these met approved sampling standards.

The content areas represented in the TIMSS mathematics test at Grade 8 were Fractions & Number Sense, Geometry, Algebra, Data Representation, Analysis & Probability, Measurement, and Proportionality. The test was designed to reflect the common or essential elements of the national mathematics curricula in participating countries. It included both multiple-choice (75%) and short-answer (25%) questions.

Irish students in Second year achieved a mean overall score (527) that did not differ significantly from the OECD country average of 526 (OECD, 1997).[11] Ireland ranked 12th among the 25 countries (both OECD and non-OECD) that met the sampling requirements (Beaton et al., 1996a). Among these, countries with higher mean scores than Ireland included Korea, Japan, and the Czech Republic. Countries with mean scores not significantly different from Ireland's included Canada, Hungary, New Zealand, and Sweden. Those with significantly lower mean scores included Norway, Iceland, Spain, and Portugal. Just 9% of students in Ireland had scores at or above the international 90th percentile. This compares

---

[9] Four of the five studies – IAEP I, IAEP II, TIMSS 1995 and PISA 2000 – also included an assessment of science.

[10] IALS also assessed the ability of adults to understand prose and documents.

[11] It may be noted that the average of all participating countries was somewhat lower at 513, but this includes countries that did not meet sampling requirements.

unfavourably with Singapore (45%), Korea (34%), Japan (32%), Hong Kong (27%), and the Czech Republic (18%).

Irish students in Second year achieved mean percent correct scores that were significantly higher than the corresponding international averages (in the 25-country group) in four mathematical content areas – Fractions & Number Sense, Data Representation, Analysis & Probability, and Proportionality (Beaton et al., 1996a). They achieved scores that were not significantly different from the corresponding international country averages in two areas – Measurement and Algebra. In one area, Geometry, Irish students achieved a mean score that was significantly lower than the corresponding international average**.**

In a comparison of the performance of students in Grades 7 and 8 in TIMSS, Irish students in Grade 8 (Second year) had a mean score that was over one-quarter of a standard deviation higher than the mean score of Grade 7 (First year) students (Beaton et al., 1996a). This difference was somewhat smaller than that found in other countries with mean scores similar to Ireland's in Grade 7. Difference scores in Hungary, the Russian Federation, the Slovak Republic, and Switzerland all exceeded one-third of a standard deviation, indicating greater relative progress than in Ireland.

PISA Mathematics (2000)

Mathematics was a minor domain in PISA 2000. Two over-arching ideas (sub-domains) were assessed – Growth & Change (called Change & Relationships in PISA 2003), and Space & Shape. However, performance was reported on a single mathematics scale only. The mean score of Irish students on this scale (502.9) did not differ significantly from the OECD country average (500.0) (Shiel et al., 2001). Ireland ranked 15th of 27 OECD countries[12], with 13 countries achieving mean scores that were significantly higher than Ireland's. These included Japan, Korea, the United Kingdom, and France. Japan, the highest scoring country, had a mean score that was over one half of a standard deviation above the mean for Ireland.

Among the countries with mean scores that did not differ significantly from Ireland's were Sweden, the Czech Republic, the United States, and Germany. Countries with significantly lower scores included Hungary, Spain, Poland, Portugal, Greece and Mexico. Only two countries (Finland and Mexico) had standard deviations that were lower than Ireland's, indicating that the distribution of achievement was narrower in Ireland than in most countries.

Irish students at the 10th percentile in PISA 2000 mathematics achieved a score that was significantly higher than the OECD country average at that marker (394.4 compared with 366.8). This ranked Ireland 14th among OECD countries on this measure. At the 90th percentile, Irish students ranked 20th and achieved a mean score that was lower than the average for OECD countries at that marker (606.2 compared with 624.8). This indicates a relatively poor performance on the part of higher-achieving students in Ireland. Indeed, Ireland had the lowest score at the 90th percentile among the countries with mean scores in the same range. On the other hand, among OECD countries, Ireland had the second smallest difference between the 25th and 75th percentiles, again indicating comparatively greater equity in learning outcomes.

An additional 11 countries administered the PISA 2000 instruments in 2002 (the 'PISA Plus' countries). When the results of all countries that completed the first cycle of PISA were combined, two additional non-OECD countries, Hong Kong-China and Liechtenstein, were found to rank higher than Ireland in mathematics (OECD/UNESCO-UIS, 2003). Of

---

[12] Note that although the Netherlands participated, achievement outcomes were not published as response rates were too low to ensure reliability.

these, only Hong Kong-China's mean score was statistically significantly higher than Ireland's.

International Assessments of Reading Literacy

As indicated in Table 2.1, Ireland has participated in three international studies of reading literacy since 1990. The results of the IEA Reading Literacy Survey may be found in Elley (1994) and Martin and Morgan (1994), while the outcomes of the IALS are described in detail in Morgan, Hickey and Kellaghan (1997) and in a number of OECD publications (e.g., OECD, 1997; OECD/Statistics Canada, 2000). In this subsection, only performance on PISA 2000 is reviewed.

PISA Reading Literacy (2000)

Reading literacy was the major domain in PISA 2000 and students were assessed on a range of texts that were Continuous (e.g., narratives, descriptions, expositions) and Non-continuous (e.g., charts, diagrams, maps, forms and tables). Performance was reported in terms of mean scores on an overall (combined) scale, on two text content/structure subscales – Continuous and Non-continuous – and on three process subscales – Retrieving information in texts, Interpreting information in texts, and Reflecting on and Evaluating the content and form of texts. Performance on the combined scale and subscales was also reported with reference to proficiency levels and scores at key markers (e.g., the 10th and 90th percentiles).

The mean performance of Irish students on the PISA combined reading literacy scale (526.7) was significantly higher than the OECD country average (500.0). Ireland ranked 5th of 27 OECD countries, with just one country (Finland) achieving a mean score which was significantly higher than Ireland. The OECD countries with mean scores that did not differ significantly from Ireland's were Australia, Canada, Japan, Korea, and Sweden. Countries with significantly lower mean scores included France, Germany, Norway, and Switzerland (Shiel et al., 2001).

Ireland ranked 4th on the Continuous texts scale, with a mean that was 27 points (one-quarter of a standard deviation) higher than the OECD country average (501). On the Non-continuous texts scale, Irish students ranked 6th, with a mean score that was 30 points (one-third of a standard deviation) higher than the OECD country average. Again, only students in Finland achieved significantly higher mean scores than Irish students on these subscales (Kirsch et al., 2002).

The mean scores of Irish students on the Retrieve and Interpret scales were about the same as on the test as a whole. Again, only students in Finland achieved significantly higher mean scores. Ireland ranked third on the Reflect/Evaluate subscale, with a mean score that did not differ significantly from Canada, the highest scoring country on the subscale. Ireland's mean scores on the three process subscales were all significantly higher than the corresponding OECD country average scores.

Five proficiency levels were identified on the combined reading literacy scale and on each of the reading literacy subscales. An additional category, 'below Level 1', was added to accommodate students whose performance did not meet the criteria for inclusion at Level 1 (the lowest proficiency level which PISA was designed to measure). In Ireland, 11.0% of students were at Level 1 or below; 17.9% at Level 2; 29.7% at Level 3; 27.1% at Level 4; and 14.2% at Level 5. The proportions of Irish students represented at each level on the Continuous texts and Non-continuous texts subscales, and on the Retrieve and Interpret process scales, were broadly similar to the percentages on the combined reading literacy scale. Performance on the Reflect/Evaluate process subscale was marginally better, with 44.0% of students achieving Levels 4 and 5.

In Ireland, the score of students at the 10th percentile on the combined reading literacy scale was 401.3. Students at the 10th percentile in just four OECD countries achieved higher scores – Korea (432.8), Finland (429.0), Canada (409.9) and Japan (407.1). Students at the 90th percentile in Ireland achieved a score of 641.1. This was exceeded by students in five countries – New Zealand (660.0), Australia (655.6), Finland (653.6), Canada (651.8), and the UK (650.7).

In 2001, eleven additional non-OECD countries participated in a second administration of the assessment. When the full pool of countries that completed the PISA 2000 assessment of reading literacy according to OECD standards between 2000 and 2002 is considered, Ireland's overall ranking is still 5th. One additional country, Hong Kong-China, performed at about the same level as Ireland, with a ranking of 6th (OECD/UNESCO-UIS, 2003).

International Assessments of Science

As indicated in Table 2.1, Ireland participated in three international studies of science between 1989 and 2000. The findings of the First International Assessment of Educational Progress may be found in Lapointe, Mead and Phillips (1989); those of the Second International Assessment of Educational Progress in Lapointe, Askew and Mead (1992), Martin, Hickey and Murchan (1992), and OECD (1993), and those of TIMSS 1995 in Beaton, Martin, Mullis, Gonzalez, Smith & Kelly (1996b) and OECD (1997). This section looks in detail at performance in science in PISA 2000.

PISA Science (2000)

Science was a minor domain in PISA 2000 and sought to measure students' ability to apply a range of scientific processes including recognising questions, identifying evidence/data, and drawing and evaluating conclusions. While some of the content areas outlined in the framework were well represented (e.g., Atmospheric Change, Earth & Universe, Energy Transfer, and Ecosystems), others (e.g., Biodiversity, Chemical & Physical Change, and Physiological Change) were not. Since it was a minor assessment domain, fewer items were used to assess science than reading (the major domain).

The mean score of Irish students on the PISA science scale (513.4) was significantly higher than the OECD country average (500.0). Ireland ranked 9th of 27 OECD countries. Six countries, including the United Kingdom, Korea, and Japan, achieved significantly higher mean scores than Ireland. Mean scores in eight countries, including Austria and Sweden, did not differ significantly from Ireland's. Thus, Ireland did comparatively better on science than on mathematics, but relatively less well than on reading literacy.

Irish students at the 10th percentile in science achieved a score that was significantly higher than the OECD country average at that marker (394.4 compared with 368.5). This ranked Ireland 7th among OECD countries. At the 90th percentile, Irish students ranked 10th and achieved a mean score that was not significantly different from the OECD average score at that marker (630.2 compared with 626.9). In Ireland, the difference between scores at the 25th and 75th percentiles is 128.8 points – some 12 points lower than the OECD country average. Again, this can be interpreted as indicating a more equitable spread of achievement in Ireland than in several other countries.

When the full pool of countries that completed the PISA 2000 assessment of science according to OECD standards between 2000 and 2002 is considered, Ireland's overall ranking in science is 10th of 41 countries. One additional country, Hong Kong-China, performed at a significantly higher level than Ireland, with an overall ranking of 3rd (OECD/UNESCO-UIS, 2003).

Comparing Performance on Mathematics, Reading and Science in PISA 2000

A relevant issue in interpreting the outcomes of PISA 2000 is the extent to which performance on the three assessment domains was inter-related. In general, countries with high achievement in one domain had high achievement in all three. For example, Australia, Canada, Finland, Japan, Korea, New Zealand, and the UK all had mean scores that were significantly above the OECD average in the three assessment domains. However, Ireland was an exception to this pattern since mean scores for reading and science were significantly higher than the corresponding OECD country average scores, but the mean score for mathematics did not differ significantly. Belgium and Iceland achieved mean scores that were above the OECD country average in reading and mathematics, but not significantly different from the OECD country average in science.

Despite Ireland's somewhat uneven performance profile, there is evidence for a strong association across assessment domains when one considers correlations between scores. The correlation between reading literacy and mathematics scores for Irish students is .82, while that between reading literacy and science is .90. A correlation of .83 was found between students' scores on science and mathematics.

VARIABLES ASSOCIATED WITH ACHIEVEMENT IN MATHEMATICS, READING, AND SCIENCE

The first part of this section describes, for students in Ireland, the relationships between a range of background variables and achievement in PISA 2000 and in earlier international studies. Variables are categorised according to whether they relate primarily to the student (e.g., background, home educational climate) or to the school (e.g., school type, disadvantaged status). The second part considers the outcomes of a series of multilevel models based on the PISA 2000 data for Ireland.

Student-level Variables and Achievement

Gender

In Ireland, male students in PISA 2000 achieved a mean mathematics score that was significantly higher (by one-sixth of a standard deviation) than the mean score for females. In contrast, on reading literacy, female students achieved a mean score that was significantly greater (by three-tenths of a standard deviation) than the mean score of male students. The difference in mean scores in favour of female students on PISA science (just over one-twentieth of a standard deviation) is not statistically significant. In reading, the difference in achievement in favour of female students is smaller on Non-continuous texts (over one-quarter of a standard deviation) than on Continuous texts (over one-third of a standard deviation). The significant gender difference in favour of female students in PISA reading literacy was not unexpected since, in the IEA Reading Literacy Study in 1991, female students (age 14) in Ireland outperformed their male counterparts by about one-quarter of a standard deviation. In mathematics, however, the significant difference in favour of males was somewhat unexpected, given that no overall difference between genders was observed in Grades 4 or 8 in TIMSS 1995, or in Fourth class in a national assessment of mathematics in 1999 (Shiel & Kelly, 2001). In the earlier IAEP II study, Irish 13-year old male students had achieved a mean score in mathematics that was significantly higher than that of female students in the same age range (Martin et al, 1992).

The gender difference in favour of female students on PISA 2000 reading literacy in Ireland was consistent with the stronger performance of females on the Junior Certificate Examination in English in 1999 (see Shiel et al., 2001), where more female students sat the Higher-level paper and achieved more A and B grades at Higher, Ordinary and Foundation levels. Although males outperformed females in PISA 2000 mathematics, equal proportions of males and females sat the Higher-level Junior Certificate Examination in Mathematics in 1999 and similar proportions achieved A and B grades at that level. While no difference between males and females was observed for PISA 2000 science, slightly more females than males took Higher-level science in the 1999 Junior Certificate Examination, and a greater proportion of females than males achieved A and B grades at both Higher and Ordinary levels.

Socioeconomic Status

In PISA 2000, students were asked to indicate their mother's and father's main occupation, and what each parent did in those occupations. Their responses were categorised using the International Socioeconomic Index (ISEI) and the highest level achieved by either parent was taken as a measure of student socioeconomic status. Correlations between SES and achievement are moderate at .29 for mathematics, .31 for reading literacy, and .24 for science for students in Ireland. For each content domain, high-SES[13] students outperformed medium-SES students, who, in turn, outperformed low-SES students. In mathematics, for example, these differences were of the order of one-fifth and one-third of a standard deviation respectively.

Parental Education

Parental education was coded using the International Standard Classification of Education (ISCED) scale, and the highest ISCED level of either parent was recorded. Correlations between this measure and achievement for students in Ireland were in the moderate range (.21 to .24) for the three domains. In each PISA 2000 assessment domain, the mean score of students with at least one parent who had completed a third-level course was significantly higher than the mean score of students with at least one parent whose highest level was upper-secondary. The largest difference – three-tenths of a standard deviation – was observed for science. Conversely, in each domain, the mean score of students of parents with no education or primary education only was significantly lower than the mean score of students of parents with upper secondary education. The largest difference between these subgroups – over two-fifths of a standard deviation – was in mathematics.

Family Structure

In PISA 2000, approximately 13% of students taking each assessment domain reported living in a lone-parent household. In all three PISA domains, these students achieved mean scores that were about one-quarter of a standard deviation lower than the scores of students living in a dual-parent household. Correlations between number of siblings and achievement are negative and weak, yet statistically significant: –.15 for science, –.11 for mathematics, and –.12 for reading literacy. The differences in mean scores in each domain between students with four or more siblings and students with two are statistically significant, with the latter group achieving scores that are between one-quarter of a standard deviation (mathematics) and one-third (reading literacy and science) higher. These data indicate that students in larger families are more likely to achieve lower scores than students in smaller families.

---

[13] In these and subsequent references to 'high', 'medium' and 'low', 'high' refers to students in the top third of a distribution (e.g., SES), 'medium' to students in the middle third, and 'low' to students in the bottom third.

Home Educational Climate Variables

Parental engagement. A composite variable, parental engagement, was constructed using students' responses to questions about the frequency with which their parents engaged with them in discussing politics or social issues, discussing books, films or television programmes, and listening to classical music. Correlations with achievement are weak but statistically significant, ranging from .10 (mathematics) to .19 (reading literacy). Although statistically significant achievement differences between students with high and medium levels of parental engagement were observed for reading literacy and science, the difference in mathematics between students in these groups is not statistically significant. In all three domains, students with medium levels of parental engagement outperformed those with low levels. The difference is somewhat greater for reading literacy (over one-quarter of a standard deviation) than for mathematics (about one-fifth of a standard deviation).

Books in the home. Students in PISA 2000 were asked to indicate the number of books (excluding magazines) at home. This measure may provide a proxy for home educational climate (through, for example, interactions between parents and students that feature books). Almost 10% of students reported having 10 or fewer books at home, while almost one-quarter had more than 250 books. The correlations between number of books in the home (where the categorical variable was treated as continuous) are all in the moderate range: .32 for mathematics and science, and .33 for reading. Students with no books in the home had significantly lower mean scores in each domain than students who had between 51 and 100 books. In mathematics, the difference between these groups is over two-fifths of a standard deviation, while in reading literacy and science it is over one-half. The number of books in a student's home was also related to mathematics and science achievement in TIMSS 1995 (Beaton et al., 1996a, 1996b), and to reading achievement in the 1991 IEA Reading Literacy Survey (Martin & Morgan, 1994).

Student as Learner

Absence from school. Students in PISA 2000 were asked to indicate the number of days on which they were absent from school in the two weeks prior to the PISA 2000 assessment, using a 3-point scale ('none', 'one or two', 'three or more'; reasons for absences were unknown). Of students who responded to this question, 57% indicated that they had not missed any days, while just under 9% said that they had missed three or more days. Whereas significant differences in performance in reading literacy and science (one-sixth of a standard deviation in each domain) were observed between students with full attendance records and those absent for one or two days, the corresponding difference for mathematics (just over one-eighth of a standard deviation) is not statistically significant.

Homework and study. Students in PISA 2000 were asked to indicate how often they completed homework on time, across all school subjects, on a 4-point scale ranging from 'never' to 'always'. Almost 5% indicated that they never completed homework on time, while just over 21% said that they always did. No differences in mean achievement scores were found for any of the PISA 2000 assessment domains between students who 'always' completed their homework on time and those who did so 'most times'. On the other hand, in all three domains, students who mostly completed their homework on time had mean achievement scores that were significantly higher (by between one-quarter and one-third of a standard deviation) than students who sometimes completed their homework on time.

Current grade level. Perhaps reflecting the fact that PISA uses an age-based sample, performance differences in PISA 2000 were observed between Irish students who were enrolled in Second, Third, Fourth (Transition), and Fifth years in the three assessment domains.[14] In mathematics, the mean difference in achievement favouring Third year students over students in Second year is 86.3 points (just over one standard deviation), while the corresponding difference for reading literacy is 106.2 points (over one and one-tenth standard deviations). It should be noted, however, that just over 3% of students were enrolled in Second year, while over 60% were enrolled in Third year. Students in Fourth (Transition) and Fifth years outperformed students in Third year in all three domains. In mathematics, the difference between Fourth year students and Third years was greater, by over one-fifth of a standard deviation, than the difference between Fifth and Third year students, perhaps reflecting differences in the age composition of Fourth (Transition) and Fifth year groups, and the curricula they experienced in relation to what is assessed by PISA.

Attitude to reading and frequency of leisure reading. A weighted composite variable, based on the responses of students to nine items designed to tap into attitude to reading, was developed in PISA 2000. The correlation between the resulting variable and PISA combined reading literacy, .43, is moderate to strong. Moreover, students in the top third of the distribution of attitude scores (those with a strong positive attitude to reading) achieved a mean score on combined reading that is one standard deviation higher than the mean score of students in the bottom third.

Students in PISA 2000 who reported engaging in leisure reading for up to 30 minutes a day achieved a mean score on combined reading literacy that is almost one-half of a standard deviation higher than the mean score of students reporting that they did not engage in any leisure reading. Similarly, those reading for 30 to 60 minutes a day achieved a score that is almost one-fifth of a standard deviation higher than the mean score of students who read for up to 30 minutes. The correlation between reading for enjoyment and performance on combined reading literacy is .26.

Self-regulated learning. Data were gathered in some countries in PISA 2000, including Ireland, on variables relating to self-regulated learning. These included use of control strategies, effort and persistence, use of memorisation, use of elaboration, self-efficacy, and instrumental motivation. Among the aspects of self-regulated learning that are most strongly correlated with performance in mathematics are mathematical self-concept (.24), control expectations (.23), and students' perceptions of their efficacy as a learner (.26). Variables with significant associations with combined reading literacy include academic self-concept (.29) and self-efficacy (.24). Variables most strongly correlated with science are academic self-concept (.29), self-efficacy (.25), and control expectations (.22). The self-regulated learning variables also correlated strongly with one another. For example, the correlation between control strategies and memorisation (both aspects of learning strategies) is .67, while that between control strategies and effort and persistence (a motivational preference) is .74. Hence, there is a danger in attributing too much importance to any single variable associated with self-regulated learning on the basis of a moderate correlation with achievement. A number of self-regulated learning variables had weak but statistically significant correlations with reading literacy, including use of elaboration strategies (.07) and use of memorisation strategies (.07).

[14] Eligible students are those born in 1987, so about one-quarter of students are actually 16 years old at the time of the PISA survey. The term '15-year-old' is used for the sake of convenience.

School Characteristics and Achievement

This section looks at school-level variables associated with achievement in PISA 2000, including school type (whether secondary, community/comprehensive, or vocational), and school designated disadvantaged status. A small number of school characteristics are based on data provided by students, which have been aggregated to the school level (e.g., negative disciplinary climate, see below).

School Type

Students in PISA 2000 attending secondary schools achieved a significantly higher overall mean score (by 21.3 points, or just over one-fifth of a standard deviation) than students in community/comprehensive schools. In mathematics and science, differences between the mean scores of students attending these school types were not statistically significant. Differences in achievement between students attending community/ comprehensive and vocational schools are significant in all three assessment domains, ranging from 38.2 points (two-fifths of a standard deviation) in reading to 23.2 points (one-quarter) in science.

School Designated Disadvantaged Status

Schools can be classified according to whether they are in the Department of Education and Science’s Disadvantaged Area Schools Scheme or not.[15] One quarter of students in PISA 2000 attended designated schools. The mean score of these students in reading literacy was 48.8 points (over one half of a standard deviation) lower than that of students attending non-designated schools. Large differences in favour of students attending non-designated schools were also observed in mathematics (38.0 points, over two-fifths of a standard deviation) and science (48.2 points, over one-half of a standard deviation).

Negative Disciplinary Climate in English Classes

In PISA 2000, students in schools with high negative disciplinary climate in English classes (as reported by the students themselves) achieved a significantly lower mean score in science (by 18.1 points, or one-fifth of a standard deviation) than students attending schools with average (medium) disciplinary climate. Mean score differences between these groups for reading literacy and mathematics were not statistically significant. Moreover, differences between students in schools with average and low negative disciplinary climate were not statistically significant in any of the PISA domains.

STATISTICAL MODELLING OF ACHIEVEMENT IN PISA 2000

In PISA 2000 in Ireland, 11.4% of the variation in achievement in mathematics, 14.1% in science, and 17.8% in reading was attributed to differences in achievement between schools. The remainder of the variance was attributed to differences within schools (i.e., between classes and students). In mathematics, only four countries (Iceland, Sweden, Norway and Finland) had lower between-school variation. The same pattern was observed in reading and science. Many of the background characteristics described above which are associated with achievement are themselves interrelated. To take the simultaneous contributions of school- and student-level variables into account, hierarchical linear models were developed for PISA 2000 reading literacy, mathematics and science in Ireland. These models also allow the quantification of explained and unexplained variance in achievement, both within and between schools.

[15] Schools in the scheme are known as ‘designated’ or ‘disadvantaged’ schools. Such schools benefit from a range of supports including additional resources (personnel and financial).

A Model of Performance in Mathematics

The model for mathematics presented in the initial national report on PISA 2000 explained 78.8% of between-school variance in achievement, and 31.9% of within-school variance (i.e., variance between classes and students). Two school-level variables were included in the final model – school type (whether secondary, community/ comprehensive, or vocational) and designated disadvantaged status. Student-level variables in the final model included gender, SES, parental education, lone-parent status, number of siblings, early school leaving intent (a measure of the likelihood that a student would leave school before the end of Senior Cycle), frequency of completion of homework on time, and current grade level. The final model also included an interaction between gender and lone-parent status. Although lone-parent status was not associated with a reduction in the achievement of male students, females living in a lone-parent family had an expected score that was one-fifth of a standard deviation lower than that of females who did not live in a lone-parent family. The model also indicated that, at least for mathematics, both student home background (based on parental occupation) and parental education make significant contributions to achievement. The index of books in the home was also found to be associated with mathematics performance, with the difference in expected outcomes between the highest (more than 500 books) and lowest (no books) points on the index exceeding one standard deviation.

Models of Performance in Reading Literacy

In the national report on PISA 2000, a model of reading literacy that explained 77.8% of between-school variance and 44.2% of within-school variance (i.e., between classes and students) was presented. The inclusion of two variables specific to reading habits and attitudes (frequency of leisure reading and attitude to reading) may account for the higher proportion of within-school variance that was explained by the model, compared to the model for mathematics. At the school level, three variables were included: school type, school designated status, and negative disciplinary climate. Student-level variables in the model included socioeconomic status, current grade level and early school-leaving intent. The model also included an interaction between gender and the number of books in the home. At the lowest levels (no books, and 1 to 10 books), male students had higher expected scores than females, while the reverse was the case at all categories from 51 to 100 books upwards. This indicates that the effect of higher levels of books in the home is stronger for females. Finally, expected reading scores of students with four or more siblings were lower than those of students with two or fewer siblings.

A follow-up (expanded) model sought to examine the effects on achievement of a number of self-regulated learning variables (Sofroniou, Shiel & Cosgrove, 2002). The expanded model explained 78.2% of variance at the school level, and 47.1% at the student/class level. Like the original model, the expanded model included the school-level variables school type, negative disciplinary climate, and school designated disadvantaged status. Also included were student socioeconomic status, number of siblings, and the frequency of completing homework on time. The expanded model included interactions between attitude to reading and number of books in the home and between frequency of absence from school and early school-leaving intent. No main effects or interactions for gender were present, indicating that gender differences were explained by other variables in the model (though the precise combination of variables that explained those differences is not known).

Although the expanded model included a number of self-regulated learning variables (academic self-concept, instrumental motivation, competitive learning, co-operative learning), their fitted contributions to achievement were modest. Moreover, variables that had been considered for inclusion, such as use of control strategies and self-efficacy for learning, were dropped during model development, indicating that their effects were accounted for by other variables.

A Model of Performance in Science

The model of science described in the national report on PISA 2000 explained 74.5% of the variance between schools, and 34.1% of the variance within schools (between classes and students). School type, school designated disadvantaged status, and negative disciplinary climate (albeit in English classes) were again included as school-level variables. The model also included an interaction between student gender and index of books in the home similar to that observed for reading. An additional student variable in the final model is of particular interest. Students who had studied science at Junior Cycle level had an estimated score that was almost one-half of a standard deviation higher than that of students who had not studied science.

THE PISA 2000 TEST-CURRICULUM RATING PROJECT

Since countries vary with respect to curricular content and delivery, international assessments of mathematics and science have usually included measures of curriculum coverage (i.e., an indication of the relatedness of each country's syllabus to the assessment). The model used in the TIMSS studies, which builds on earlier IEA studies (e.g., Travers & Westbury, 1989), is a useful way of conceptualising curriculum. In this model, curriculum is considered in three components – the *intended curriculum* (what a particular education system expects students to learn and how the system should be organised to facilitate this learning, as indicated, for example, by instructional goals), *the implemented curriculum* (what is taught in classrooms and how it is taught, as indicated, for example, in textbooks and teachers' reports), and the *attained curriculum* (what students have learned, evidenced in outcomes on formal assessments or other outcome measures) (see Mullis et al., 2003, pp. 3-5).

Evidence of 'opportunity to learn' was gathered in TIMSS by asking curriculum experts to indicate whether or not the content of each item appeared on the national curriculum (a measure of the intended curriculum) (see, for example, Table B.1 in Beaton et al., 1996a).[16] Since TIMSS sought to ensure that test items were appropriate for the students of all participating countries and reflected their current curricula (e.g., Beaton et al., 1996a, b), the implementation of the opportunity-to-learn measure followed naturally from this. PISA, on the other hand, sought to assess not only school-based learning but also the wider knowledge and skills needed by adults in society and thus does not include an opportunity-to-learn measure.

However, given the apparent differences in the content and approach between the Irish Junior Certificate syllabi/examinations and PISA 2000, an analysis of performance on the three assessment domains was conducted by the Educational Research Centre to examine the relative strengths and weaknesses displayed by Irish students in terms of what they might reasonably be expected to learn in school.

Framework for the Test-Curriculum Rating Project

The aim of the test-curriculum rating project was to develop a set of rating scales which are capable of capturing the extent and type of similarities and differences between

[16] Schmidt et al. (1997) extended analyses of test content in TIMSS 1995 to a detailed examination of textbooks and curriculum documents on a country-by-country basis, although much of this work is descriptive and does not establish links between curriculum-relatedness and achievement in individual countries.

PISA test items in reading, mathematics and science, and the types of questions students in Third year (Grade 9) of the Junior Cycle are exposed to[17], based on an examination of the intended curriculum at each level of the syllabus. The intended curriculum was defined in the same way as in TIMSS: instruction and learning goals in mathematics as defined at the system level (see Beaton et al., 1996a, pp. A1-A2), and encompassed in Irish syllabus documents, teacher guidelines, and Junior Certificate Examination papers.

Three points should be noted in interpreting outcomes of the test-curriculum rating project, which compared the intended Junior Certificate Examination syllabus to PISA. First, the syllabus may not be implemented as intended at all times. Second, the test-curriculum rating project did not take into account the likelihood that numerous factors, other than curriculum intent and the manner in which it is implemented, affect student achievements (the attained curriculum). Third, several aspects of the Irish syllabi are not assessed by PISA.

Procedure and Definition of Rating Scales

For each PISA domain, three qualified teachers with extensive experience in their subject area in Ireland (ranging from curriculum development and examination setting to teaching) carried out the curriculum rating exercise. The rating scales are based on a framework that differentiates levels of expected familiarity with an item cross-tabulated with *item aspect* and *syllabus level.* In mathematics, the framework comprises a 3 x 3 matrix, containing three item aspects: mathematics concept, context of application of mathematics concept, and mathematics item format (Table 2.2).

Table 2.2. Framework for the 2000 Test-Curriculum Rating Project: Mathematics Items

	Junior Certificate Level		
Aspect	*Higher*	*Ordinary*	*Foundation*
Concept: How familiar would you expect the typical Third year student to be with the specific mathematical concept(s) underlying this item?	Not/Somewhat/Very Familiar	Not/Somewhat/Very Familiar	Not/Somewhat/Very Familiar
Context/Application: How familiar would you expect the typical Third year student to be with the application of the specific mathematical concept(s) underlying this item in the type of context suggested by the item and stimulus text?	Not/Somewhat/Very Familiar	Not/Somewhat/Very Familiar	Not/Somewhat/Very Familiar
Format: How familiar would you expect the typical Third year student to be with the application of the specific mathematical concept(s) underlying this item in the type of format suggested by the item and stimulus text?	Not/Somewhat/Very Familiar	Not/Somewhat/Very Familiar	Not/Somewhat/Very Familiar

Using this framework, each mathematics item was rated for each aspect in terms of its expected familiarity for a typical student at the three syllabus levels. Hence, each item received nine ratings. Ratings on individual aspects ranged from 1 (‘not familiar’) to 3 (‘very familiar’). Initially, raters rated items independently, and items on which there was a lack of consensus were flagged. Consensus was achieved at rater meetings, and flagged items were assigned a modal rating.

Rating Outcomes for Mathematics

The expected familiarity of students with the mathematical concepts tapped by PISA 2000 items was fairly evenly spread across the three scale points (not/somewhat/very

[17] Third year (Grade 9) was the modal grade for PISA 2000 students: 62% were in Third year at the time of the assessment; just 3% were in Second year, and the remainder were in Fourth (transition) or Fifth year.

familiar) (Shiel et al., 2001).[18] Comparing Higher and Foundation levels, the percentage of items rated as 'very familiar' drops from 28.1% to 9.4% (Table 2.3). In terms of context of application, no PISA 2000 mathematics item was rated as 'very familiar' (at any syllabus level) and, in fact, between 71.9% and 81.3% of all items were rated as 'not familiar' on this scale, depending on syllabus level. Familiarity with item format was also low, with no items rated as 'very familiar' for any syllabus level and between 53.1% and 78.1% rated 'not familiar', depending on the syllabus level in question.

Taking into account both the mathematics syllabus level studied by each student for the Junior Certificate and the particular set of items the student attempted during the PISA 2000 assessment[19], the average expected familiarity of each student with those items that s/he attempted was computed for each of the three aspects (concept, context of application, and item format). These student-level familiarity ratings were then related to students' performance on PISA 2000 mathematics. The correlation between achievement and concept familiarity (.48) was higher than that between achievement and context/application (.23) or between achievement and item format (.20). This indicates that familiarity with the mathematics concepts assessed in PISA predicted achievement better than familiarity with the contexts in which the concept was applied or the response formats of the items.

Table 2.3. Percentages of Ratings Assigned to PISA 2000 Mathematics Items, by Scale and Syllabus Level (N items = 32)

	Not Familiar	*Somewhat Familiar*	*Very Familiar*	*Total*
Concept				
Higher	31.3	40.6	28.1	100.0
Ordinary	34.4	46.9	18.8	100.0
Foundation	53.1	37.5	9.4	100.0
Context/Application				
Higher	71.9	28.1	0.0	100.0
Ordinary	75.0	25.0	0.0	100.0
Foundation	81.3	18.8	0.0	100.0
Format				
Higher	53.1	46.9	0.0	100.0
Ordinary	78.1	21.9	0.0	100.0
Foundation	71.9	28.1	0.0	100.0

Source: Shiel et al. (2001), Table 6.19.

Rating Outcomes for Reading Literacy and Science

In reading, items in PISA 2000 were rated in terms of students' expected familiarity with the underlying processes, the contexts/applications in which they were expected to apply those processes (as represented by the items), and the formats of the items. At Higher and Ordinary levels, the processes underlying the PISA items were rated as very familiar or familiar in over 90% of cases. At Foundation level, 75.0% of items received these ratings. While the expected familiarity of context/application ratings dropped from 86.7% (very familiar/familiar) at Higher level, to 81.6% at Ordinary level, and 49.1% at Foundation level, this is perhaps not surprising as students taking Foundation level would not be expected to

[18] These analyses refer to the pre-2000 mathematics syllabus. A revised Junior Certificate Mathematics Syllabus, hereinafter called the Revised Junior Certificate Mathematics Syllabus, was implemented from 2000 onwards, with first examination in 2003.

[19] In PISA 2000, the assessment items were organised into a series of half-hour blocks rotated across nine test booklets. Each student attempted four such blocks. Five of the nine booklets contained at least one block of mathematics items.

read texts of the same level of complexity as students at Higher or Ordinary levels. Items rated as not familiar in terms of process and/or context tended to be associated with more complex non-continuous texts. At least 50% of PISA items were rated as being 'unfamiliar' in format at each syllabus level, indicating that, relative to the Junior Certificate Examination in English, PISA included more multiple-choice items. As with mathematics, student-level familiarity ratings were related to students' performance on reading. The correlation between achievement and familiarity with process (.55) is about the same as that between achievement and context/application (.54). The correlation between achievement and familiarity with item format is .46.

In science, items in PISA 2000 were rated in terms of the expected familiarity of students taking each syllabus level (Higher and Ordinary only)[20] with the underlying processes, concepts, contexts, and item formats. Students taking science were more familiar with the processes underlying science items (91.4% of items were rated as 'very familiar' or 'familiar' at both levels) than with the underlying concepts (51.5% at Higher level; 45.8% at Ordinary level). In the case of 80.0% of items at each level, the contexts in which students were expected to demonstrate knowledge of concepts and applications were rated 'unfamiliar', while the formats in which 42.9% of the items at each level appeared were also rated unfamiliar. Correlations between dimensions of the rating scales and performance on PISA science were weak for the process (.05), context (–.01) and format (.06) scales (and non-significant in the case of context), and in the weak to moderate range for the expected familiarity with concept scale (.19). These correlations did not include scores for students in PISA 2000 who indicated that they did not study science as a subject for the Junior Certificate.

In a separate analysis, in which each PISA item was rated in terms of its location in the Junior Certificate syllabus, 30% of items were rated belonging to basic science (the core element of the syllabus), while 23% were rated as belonging to the Earth science option. Just over 40% of PISA science items could not be located in the Junior Certificate syllabus.

CONCLUSION

The overall achievement outcomes in mathematics for Ireland in TIMSS in 1995 (Second year) and PISA 2000 (15-year-old students) are broadly similar in that, in both studies, Irish students achieved mean scores that do not differ significantly from the corresponding OECD country average scores. This is despite the fact that TIMSS is based on the curricula of participating countries to the extent that there is a 'common denominator' across mathematics curricula (Oldham, 1989), while PISA is based on the mathematical knowledge and skills that students are thought to need in their future lives as citizens and as life-long learners.

In a review of the performance of students in Ireland on PISA 2000 mathematics, Oldham (2002) noted several factors that might have impacted on performance. These included the non-standard and non-routine style of the PISA questions, which contrasted with the more context-free format used in the Junior Certificate Mathematics Examination, and the poor match between the content of the Junior Certificate syllabus and the PISA assessment framework. Oldham also noted the strong influence of modern mathematics and structuralism on mathematics education in Ireland, and expressed the concern that, if taught by teachers who are unfamiliar with the underlying rationale, the curriculum could become mechanistic (and hence, students might not acquire the flexibility needed to deal with items

[20] Refers to the 1989 Junior Certificate Science Syllabus, which was studied by all students of science who participated in PISA 2000 and PISA 2003. A new syllabus was implemented in most post-primary schools in First year in 2003, with first examination scheduled for 2006.

such as those in PISA). She further noted, however, that gaps in national curricula (relative to PISA mathematics) may well be justifiable in light of countries' own educational priorities and practices, and the fact that, for many countries with high retention rates at upper second level, topics not covered by age 15 might be covered later on.

The relatively strong performance of Irish students on PISA reading literacy in 2000 was somewhat unexpected. The 1991 IEA Reading Literacy Study had indicated that Irish 9- and 14-year olds were average compared with students in other countries (Martin & Morgan, 1994), while the outcomes of the IALS study suggested that a comparatively large proportion of Irish adults had literacy difficulties (Morgan et al., 1997; OECD, 2000). As the outcomes of the curriculum-rating activity in English suggest, the relatively good match between the Junior Certificate English Syllabus and PISA reading literacy may be one reason why students in Ireland did well.

The performance of Irish students in science in PISA 2000 was also better than expected. Earlier studies suggested that achievement relative to other countries was poor. Irish 13-year-olds in the 1991 IAEP II study ranked 9th of 10 participating OECD countries (and 14th of 15 countries), and achieved a mean score that was significantly lower than the OECD and international averages (Lapointe, Askew, & Mead, 1992; OECD, 1993). Irish students in Second year in TIMSS 1995 achieved a mean score in science that was not significantly different from the OECD and international averages; weaknesses were identified in Chemistry and Physics (Beaton et al., 1996b; OECD, 1997). While the curriculum-rating analysis for science summarised in this chapter suggests a mismatch between PISA and the Junior Certificate Science Syllabus in 2000 (with 40% of PISA items not located in the syllabus), it is also possible that PISA science may play to the strengths of Irish students, with its particularly strong focus on biology and Earth science. Finally, given the relatively strong performance of Irish students in PISA reading, the presence of complex texts in some of the PISA science items may have been to the benefit of Irish students.

In considering why Irish students did less well than students in the UK in PISA science, Cosgrove, Shiel and Kennedy (2002) noted that, while science was a compulsory subject in England and Wales at both primary and post-primary levels, it was not due to be implemented as a subject in its own right at primary level in Ireland until September 2003; and that it is not taken by about 11% of students at post-primary level (many of them attending all-girls schools) (see also Task Force on the Physical Sciences, 2002). It was also noted that there was a stronger emphasis on the assessment of science at national level in England and Wales (all students participate in a national assessment at ages 11 and 14), and a greater emphasis on implementation of the scientific method in science curricula and in scientific investigations (mini-projects in science) than was required by the Junior Certificate science syllabus in place in Ireland in 2000.

Relative to other countries, between-school variation in the achievement of students in Ireland in PISA 2000 was low. In mathematics, for example, only four countries (Iceland, Sweden, Norway and Finland) had lower between-school variation in achievement. This can be interpreted as indicating that, relative to schools in most other OECD countries, schools in Ireland are similar to one another in terms of achievement. The between-school variance estimates for Ireland are considerably lower than might be expected on the basis of earlier studies. For example, in an analysis of the data for Irish students in Grade 8 in TIMSS 1995, between-school variance estimates of 44% and 33% for mathematics and science respectively were found (see Sofroniou, Cosgrove & Shiel, 2002). However, TIMSS does not provide a good estimate of between-school variance as its sample design entailed the selection of an intact mathematics class in each school, while PISA selected a random sample of 15-year-olds across classes and grade (year) levels.

PISA 2000 revealed associations between achievement and a range of student and school variables. At the student level, these were gender, socioeconomic status (based on parental occupations), parental education, family structure, parental engagement in a student's learning, number of books in the home, and self-regulated learning attitudes and strategies. At the school level, they included school disadvantaged status (a measure of a school's socioeconomic status), school type (secondary, community/comprehensive, vocational) and school disciplinary climate. While all of these variables were associated with achievement, the multi-level models that were subsequently developed highlight the importance of considering the simultaneous contributions of such variables to achievement.

The models explained over three-quarters of between-school variance in achievement in all three assessment domains, and between 32% (mathematics) and 47% (reading) of within-school variance (i.e., between classes and students). The relatively larger proportion of within-school variance explained by the reading literacy models (both initial and expanded) may be due to the presence of a few additional variables that are specific to reading, including attitude to reading and frequency of leisure reading.

A notable feature of the models concerns the contributions of both school-level socioeconomic status (represented by school designated disadvantaged status) and student-level socioeconomic status (based on parental occupations), after adjusting for the effects of other variables. The models also include variables associated with home educational processes, including parent engagement in learning and completion of homework on time. Although strongly correlated with achievement, the effect of the number of books in the home is difficult to interpret since the precise ways in which it impacts on achievement are unclear. The 'expanded' model of reading literacy is notable to the extent that the inclusion of a number of variables associated with self-regulated learning did not account for a sizeable increase in the proportions of between- and within-school variance explained by the initial reading literacy model. This suggests that variables such as academic self-concept, instrumental motivation and preference for particular learning styles should be looked at critically in terms of how they are operationalised and measured in cross-sectional surveys such as PISA, and also whether they may be better considered as joint outcomes alongside achievement.

In addition to models summarised in this chapter, which were designed to explain performance on PISA, additional models have been developed to explain performance on the Junior Certificate Examination in English (Sofroniou, Shiel & Cosgrove, 2000), and in mathematics and science (Sofroniou, Cosgrove & Shiel, 2002) using the PISA background variables. The considerable degree of similarity between the Junior Certificate and PISA models provides further confirmation of the association between many of the variables considered in this chapter and achievement in general, as well as parallels between the Junior Certificate Examination and aspects of the PISA assessments.

Finally, in analyses of links between the content of three PISA 2000 assessment domains and the corresponding Junior Certificate subjects, a moderate overlap was revealed between the concepts assessed in PISA 2000 mathematics and those covered in the Junior Certificate mathematics syllabus in place in schools in 2000. The contexts in which mathematics items were presented, and their formats, were judged to have a low level of familiarity for students in Ireland.

3

The Performance of Irish Students on PISA 2003 in an International Context

In this chapter, the performance of Irish 15-year-olds on the PISA assessments of mathematics, reading literacy, and science in 2003 is described in the context of the performance of students in other participating OECD and non-OECD (partner) countries. First, performance is described in terms of achievement on the combined mathematics scale and four mathematics subscales. Mean achievement, the distribution of achievement in terms of performance at key percentiles, and the proportion of students achieving each PISA proficiency level are examined. In the second and third sections, achievement outcomes for the minor domains of reading literacy and science are described, again looking at mean performance and the distribution of performance. In the fourth section, achievement in mathematics, reading literacy and science in PISA 2000 and PISA 2003 are compared within and across the 32 countries for which such a comparison is possible. In the fifth section, associations between the three assessment domains are considered. A summary of the main results for Ireland is presented in the conclusion.

The results for PISA problem solving are described in Chapter 7. Readers interested in gender and other correlates of achievement are referred to Chapter 4, where associations between these variables and achievement are described.

SCALING STUDENT ACHIEVEMENTS

PISA achievement data were scaled using Item Response Theory (IRT), which has the advantage of placing both student achievement and item difficulty on the same metric. IRT is compatible with the approach in PISA in which individual students attempt only a proportion of the total item set in each domain.

In 2000, the PISA achievement scales in mathematics, reading and science were each set to have an OECD mean score of 500.0 and a standard deviation of 100.0, using random samples of 500 students drawn from each participating OECD country. To be able to compare achievement in reading and science in 2000 with that in 2003, the 2003 achievement scales were anchored to the 2000 ones so that in 2003, the OECD mean may not be exactly 500.0, and the OECD standard deviation not exactly 100.0. The OECD mean for reading literacy in 2003 is actually 494.2, and the standard deviation is 100.2.

Because the PISA 2003 assessment of mathematics expanded from two areas (Space & Shape, Change & Relationships) to four (with the addition of Quantity and Uncertainty), overall mathematics scores cannot be compared for 2000 and 2003, but scores can be compared on the Space & Shape and Change & Relationships subscales. The PISA 2000 combined mathematics scale was re-scaled to these two subscales to make comparisons possible, while a new combined mathematics scale was prepared for 2003.

ACHIEVEMENT IN MATHEMATICS

In this section, the results for PISA 2003 mathematics are summarised in terms of the overall performance of students, the proportions of students achieving each of the PISA proficiency levels, and the scores of students at key markers (5th, 10th, 25th, 75th, 90th, and 95th percentiles). Results are provided for the combined mathematics scale, which is a measure of performance on the full set of PISA 2003 mathematics items, and for the four mathematics subscales, which correspond to the mathematical content areas of Space & Shape, Change & Relationships, Quantity, and Uncertainty.

Interpreting Scores on the Mathematics Scales

As indicated in Chapter 1, students were assessed on their ability to interpret schematic, numerical and textual information, and to respond to mathematics questions based on this information. Each question was categorised according to the mathematical content area or over-arching ideas that it was judged to represent. Questions were also classified according to the amount of mathematical reasoning or abstraction involved, which ranges from simple *Reproduction* of practised knowledge, facts and routine procedures, to making *Connections* across parts of the stimulus or across mathematical areas to find the solution in a less routine setting, to problems requiring *Reflection* or significant amounts of mathematical abstraction, reasoning, modelling and/or argumentation. Examples of mathematics items can be found in Appendix A, and are also discussed with reference to the Irish Junior Certificate Mathematics Examination in Chapter 6 and Appendix C. For analytic purposes, a scale was developed for combined mathematics (i.e., based on performance on all item types), and separate subscales were developed for each of the four mathematical content areas. A note on the interpretation of the achievement outcomes may be found in Inset 3.1.

Mean Scores on the Combined Mathematics Scale

Ireland achieved a mean score of 502.8 on the combined mathematics scale, which is not significantly different from the OECD country average of 500.0 (Table 3.1). The Irish mean merits a ranking of 17th among 29 OECD countries (95% confidence interval for Ireland's ranking = 15th to 18th) and a ranking of 20th among 40 OECD and partner countries (95% confidence interval for Ireland's ranking = 17th to 21st). Ireland's mean score is not significantly different from those of eight OECD countries (the Czech Republic, Denmark, France, Sweden, Austria, Germany, the Slovak Republic, and Norway), and is significantly lower than those of 10 OECD countries, including Finland, Korea, the Netherlands, Japan, Canada, Australia and New Zealand. Ten OECD countries, including Poland, Hungary, Spain, the USA and Greece, achieved significantly lower mean scores.

The table includes an indication of population coverage since not all 15-year-olds in a country are listed in the sampling frame. Some students will have left school before the time of assessment, others are below grade 7, the minimum grade included in PISA 2003, and in some countries substantial numbers of students may also commute across country borders to attend school. Since the reliability of the actual percentage enrolment values is not always good, with several countries indicating values of 100% enrolment or more, enrolment rates were categorised into three levels to give a broad indication of the degree of coverage. Lower scoring, less developed countries tend to have poorer enrolment rates and this suggests that their mean PISA scores will be upwardly biased estimates of the literacy achievements of 15-year-olds in their populations, since it is likely that substantial numbers will have left school or will not have reached Grade 7.

Countries tend to cluster on the combined mathematics scale at upper and middle rankings, while there is more variation in mean scores at the lower rankings. For example, just 17.9 score points separate the countries in the top seven positions. In contrast, 67.4 score points separate the countries in the bottom seven positions. While it can be observed that 194.4 score points – close to two standard deviations – separate the top and bottom countries (Hong Kong-China and Brazil), and 159.1 points separate the top and bottom OECD countries (Finland and Mexico), there is more variation in achievement within individual countries, as evidenced in large standard deviations in some countries (e.g., Belgium and Germany). This point is further developed in the next section, where the distribution of achievement within countries is analysed in greater detail.

Inset 3.1. Interpreting Achievement Outcomes in PISA 2003

*OECD average and OECD total***.** Some of the tables in this chapter include both *OECD average* and *OECD total* estimates. The OECD average is the mean which is obtained when each participating OECD country receives equal weight. The OECD total, on the other hand, assigns a weighting for each country based on its population of 15-year-olds, so that larger countries contribute more to the total. Given that multiple comparisons of country performance are made with reference to the OECD average, it rather than total is the statistic referred to in the text of the chapter. OECD total is included in the tables for readers interested in making comparisons on that basis.

Standard errors and confidence intervals. The statistics in this chapter are estimates of performance based on samples of students who attempted subsets of the PISA items. The standard error provides an estimate of the degree to which a statistic (such as a country mean score) may be expected to vary about the true (but unknown) population mean. If a Normal distribution is assumed, a 95% confidence interval for a mean (consisting of a region from 1.96 standard errors below the mean to 1.96 standard errors above the mean) may be constructed in such a way that, if the sampling procedure were repeated a large number of times, and the sample statistic re-computed each time, the confidence interval would be expected to contain the population estimate 95% of the time. The mean score for combined mathematics for Irish 15-year-olds in PISA is 502.3, with a standard error of 2.44. Hence, it can be stated with 95% confidence that the population mean lies in a band that extends from 497.5 to 507.1 (i.e., $502.3 \pm 1.96 \times 2.44$).

Standard deviation. The standard deviation associated with a mean score provides an indication of the dispersion of scores in a country: the smaller the value, the less dispersed the scores are. Within a given country, 68% of scores fall within plus or minus one national standard deviation of the mean score (again assuming a Normal distribution of scores). Hence, in the case of Ireland, which has a comparatively small standard deviation of 85.3 on combined mathematics (Table 3.1), 68% of Irish students' scores fall within the interval 417.0 to 587.6 (502.3 ± 85.3).

Comparing country mean scores. Comparisons were drawn between the mean achievement scores of participating OECD member and partner countries who met agreed criteria with regard to the sampling of schools and students. The comparisons, in which the differences between a country's mean score and those of the other 39 qualifying countries were examined, took into account the standard errors of measurement associated with pairs of mean scores, using a statistic called the standard error of the difference. Further, because several comparisons were made simultaneously, the critical values associated with the statistical significance of mean score differences were adjusted to more conservative levels, with reference to the number of comparisons being made, using the Bonferroni procedure for multiple comparisons. A second approach to comparing country mean scores, the Nonparametric Maximum Likelihood method, which is also used in this chapter, is described in Inset 3.2.

Omission of the United Kingdom. Readers should note that the UK did not meet the required sampling response rate standards, with an achieved initial response rate of 63% at the school level and 78% at the student level. It is not possible to quantify the impact of non-response on the estimates of achievement scores. The initial international report on PISA 2003 (OECD, 2004b) includes achievement estimates for Scotland, which did meet the sampling requirements.

Interpretation of outcomes for countries with low enrolment rates. The performance of Mexico and Turkey should be interpreted with reference to their comparatively low enrolment rates of 15-year-olds (58% and 54%, respectively). In addition, Brazil, Indonesia and Uruguay have enrolment rates of 65%, 73%, and 74%, respectively, and Liechtenstein, Thailand, and Macao-China have enrolment rates between 80% and 90%. All other participating countries have enrolment rates that equal or exceed 90% of the population of 15-year-olds, and 28 participating countries including Ireland have enrolment rates that equal or exceed 95%.

Table 3.1. Mean Achievement Scores and Standard Deviations on Combined Mathematics – OECD and Partner Countries

	Pop	Mean	(SE)	SD	(SE)	OECD Diff		Pop	Mean	(SE)	SD	(SE)	OECD Diff
Hong Kong-Ch	⊠	550.4	(4.54)	100.2	(3.01)	▲	Norway	⊠	495.2	(2.38)	92.0	(1.15)	▼
Finland	⊠	544.3	(1.87)	83.7	(1.08)	▲	Luxembourg	⊠	493.2	(0.97)	91.9	(0.95)	▼
Korea	⊠	542.2	(3.24)	92.4	(2.14)	▲	Poland	⊠	490.2	(2.50)	90.2	(1.34)	▼
Netherlands	⊠	537.8	(3.13)	92.5	(2.33)	▲	Hungary	⊠	490.0	(2.84)	93.5	(1.96)	▼
Liechtenstein	□	535.8	(4.12)	99.1	(4.43)	▲	Spain	⊠	485.1	(2.41)	88.5	(1.26)	▼
Japan	⊠	534.1	(4.02)	100.5	(2.75)	▲	*Latvia*	⊠	483.4	(3.69)	87.9	(1.66)	▼
Canada	⊠	532.5	(1.82)	87.1	(0.97)	▲	United States	⊠	482.9	(2.95)	95.2	(1.29)	▼
Belgium	⊠	529.3	(2.29)	109.9	(1.78)	▲	*Russian Fed*	⊠	468.4	(4.20)	92.3	(1.93)	▼
Macao-Ch	□	527.3	(2.89)	86.9	(2.41)	▲	Portugal	⊠	466.0	(3.40)	87.6	(1.66)	▼
Switzerland	⊠	526.6	(3.38)	98.4	(2.05)	▲	Italy	⊠	465.7	(3.08)	95.7	(1.87)	▼
Australia	⊠	524.3	(2.15)	95.4	(1.50)	▲	Greece	⊠	444.9	(3.90)	93.8	(1.76)	▼
New Zealand	⊠	523.5	(2.26)	98.3	(1.17)	▲	*Serbia & Monte.*	⊠	436.9	(3.75)	84.7	(1.55)	▼
Czech Rep	⊠	516.5	(3.55)	95.9	(1.87)	▲	Turkey	×	423.4	(6.74)	104.7	(5.34)	▼
Iceland	⊠	515.1	(1.42)	90.4	(1.21)	▲	*Uruguay*	×	422.2	(3.29)	99.7	(1.60)	▼
Denmark	⊠	514.3	(2.74)	91.3	(1.44)	▲	*Thailand*	□	417.0	(3.00)	82.0	(1.79)	▼
France	⊠	510.8	(2.50)	91.7	(1.80)	▲	Mexico	×	385.2	(3.64)	85.4	(1.85)	▼
Sweden	⊠	509.0	(2.56)	94.7	(1.79)	▲	*Indonesia*	×	360.2	(3.91)	80.5	(2.06)	▼
Austria	⊠	505.6	(3.27)	93.1	(1.67)	o	*Tunisia*	⊠	358.7	(2.54)	82.0	(1.95)	▼
Germany	⊠	503.0	(3.32)	102.6	(1.77)	o	*Brazil*	×	356.0	(4.83)	99.7	(2.95)	▼
Ireland	⊠	**502.8**	**(2.45)**	**85.3**	**(1.26)**	o	**OECD Total**		**489.0**	**(1.07)**	**103.6**	**(0.74)**	
Slovak Rep	⊠	498.2	(3.35)	93.3	(2.32)	o	**OECD Average**		**500.0**	**(0.63)**	**100.0**	**(0.43)**	

⊠	>90% of 15-year olds enrolled	■	Mean significantly higher than Ireland	▲	Above OECD average
□	75-90% of 15-year olds enrolled	□	Mean not significantly different from Ireland	o	At OECD average
×	50-75% of 15-year olds enrolled	▒	Mean significantly lower than Ireland	▼	Below OECD average

Note. OECD countries are in regular font; partner countries are in italics. SD = Standard deviation; SE = Standard error.

The column "Pop" is an indicator of the percent of the 15-year-old population enrolled in schools in each country and is based on Column 15 of Table A3.1 in OECD (2004b).

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.

When countries are grouped into distinct performance groupings for combined mathematics using the nonparametric maximum likelihood (NPML) estimation method (see Inset 3.2), six separate groups are identified (Table 3.2). Figure 3.1 shows a plot of the NPML probability distribution as well as a Normal approximation for the six masspoints with the same mean and standard deviation. The NPML empirical distribution has a negative skew with small groups of low performing countries lying to the left of the main distribution. Ten countries fall into the top grouping, which has a mean of 535.7. These include Finland, Korea, the Netherlands, and Japan, and, amongst partner countries, Hong Kong-China, Liechtenstein, and Macao-China. There is a group of four countries whose membership of the adjacent groups is uncertain, between the first and second groupings. Ireland falls into the second grouping (mean of 504.6), along with France, Sweden, Austria, Germany, and the Slovak Republic, while Denmark and Norway have a probability of belonging to this group that equals or exceeds .90. This is followed by four more countries whose group membership is uncertain. The USA, Portugal, and Italy are in the third group (mean of 475.3), together with the partner countries Latvia and the Russian Federation. The fourth grouping (mean of 428.1) comprises Greece and Turkey and partner countries Serbia and Montenegro, Uruguay, and Thailand. Mexico is alone in the fifth group (with a masspoint value of 385.0) and Indonesia, Tunisia, and Brazil form the sixth group (with a value of 358.0).

Inset 3.2. Comparing the Performance of Countries in PISA 2003: Bonferroni Adjustment Method and Nonparametric Maximum Likelihood Estimation Method

Simple rankings of countries do not take account of the standard errors of mean achievement. Tables such as Table 3.1 and the tables of multiple comparisons of mean achievement in the initial OECD report on PISA 2003 (OECD, 2004b) present *Bonferroni multiple comparisons* of 40 countries to take this error into account. The number of other countries compared to a single country gives 39 comparisons. The overall error probability or alpha-level is 0.05 per comparison. Using the Bonferroni method, the alpha-level is 0.05/39, i.e., 0.0013. This adjustment results in conservative estimates of the critical values, especially when the number of countries is large; i.e., the bands of non-significant difference are very wide. The per-comparison rate would, of course, be different when comparing smaller groups, e.g., EU countries, Nordic countries.

Another way to make these comparisons is to use *nonparametric maximum likelihood* (NPML) *estimation* to fit a form of latent class model to classify the countries into empirically distinguishable performance groupings (Atkin, Francis, & Hinde, in press). The number of groups corresponds to the maximum number of discrete masspoints required for the fitted probability distribution. Each masspoint has a mean score associated with it, as well as a proportion which is the probability of an unknown country from this population falling into that grouping. Using this method, countries are grouped if their posterior probabilities of belonging to a common masspoint are >= 0.95 (a convenient convention). Borderline values are between 0.90 and 0.95. Analyses for PISA 2000 using this method were presented by Sofroniou (May, 2004) and are available from the author upon request.

When one compares the outcomes for this method with the Bonferroni method, one can see that additional information is obtained from the data. For example, rather than three groups (significantly higher, same as, significantly lower), the data for the combined mathematics scores form six distinct groups. The method discriminates particularly well among lower-achieving countries. The fitted distribution is distinctly non-Normal, with a negative skew suggesting a 'long tail' of lower-performing countries.

Figure 3.1. Plots of the NPML Probability Distribution (left) and Corresponding Normal Approximation (right) for the Combined Mathematics Country Scores

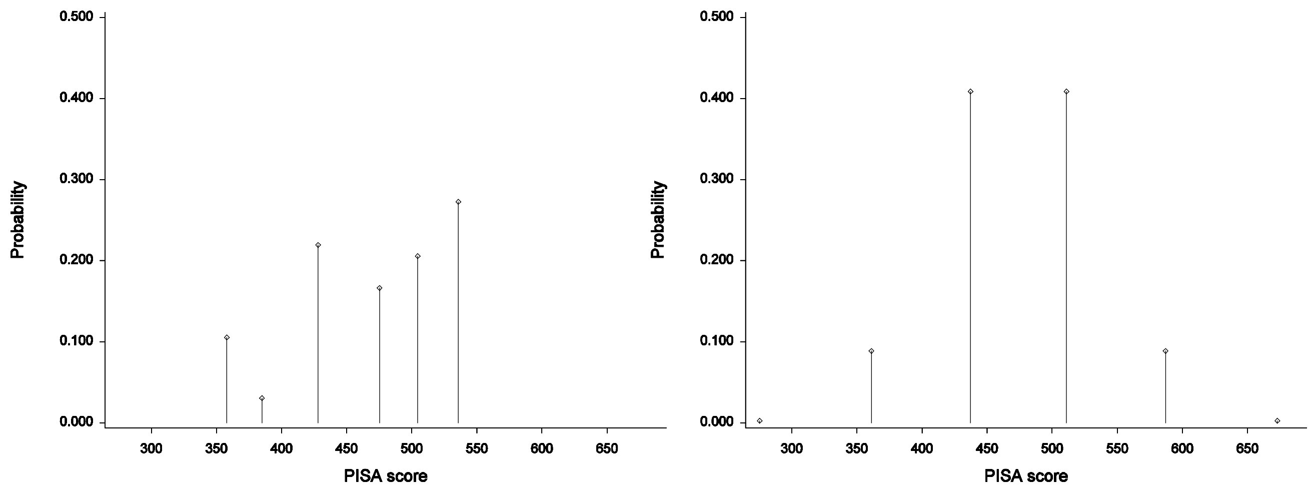


Table 3.2. *Six-point NPML Probability Distribution and Posterior Probabilities for the Combined Mathematics Scale – OECD and Partner Countries*

	Rank	*Masspoints*					
PISA score		*358.0*	*385.0*	*428.1*	*475.3*	*504.6*	*535.7*
Proportion		*0.105*	*0.03*	*0.219*	*0.166*	*0.206*	*0.273*
Hong Kong-China	1	0.00	0.00	0.00	0.00	0.00	1.00
Finland	2	0.00	0.00	0.00	0.00	0.00	1.00
Korea	3	0.00	0.00	0.00	0.00	0.00	1.00
Netherlands	4	0.00	0.00	0.00	0.00	0.00	1.00
Liechtenstein	5	0.00	0.00	0.00	0.00	0.00	1.00
Japan	6	0.00	0.00	0.00	0.00	0.00	1.00
Canada	7	0.00	0.00	0.00	0.00	0.00	1.00
Belgium	8	0.00	0.00	0.00	0.00	0.02	0.98
Macao-China	9	0.00	0.00	0.00	0.00	0.04	0.96
Switzerland	10	0.00	0.00	0.00	0.00	0.05	0.95
Australia	11	0.00	0.00	0.00	0.00	0.12	0.88
New Zealand	12	0.00	0.00	0.00	0.00	0.16	0.84
Czech Republic	13	0.00	0.00	0.00	0.00	0.78	0.22
Iceland	14	0.00	0.00	0.00	0.00	0.86	0.14
Denmark	15	0.00	0.00	0.00	0.00	0.90	0.10
France	16	0.00	0.00	0.00	0.00	0.97	0.03
Sweden	17	0.00	0.00	0.00	0.00	0.99	0.01
Austria	18	0.00	0.00	0.00	0.00	1.00	0.00
Germany	19	0.00	0.00	0.00	0.00	0.99	0.00
Ireland	**20**	**0.00**	**0.00**	**0.00**	**0.01**	**0.99**	**0.00**
Slovak Republic	21	0.00	0.00	0.00	0.03	0.97	0.00
Norway	22	0.00	0.00	0.00	0.09	0.91	0.00
Luxembourg	23	0.00	0.00	0.00	0.18	0.82	0.00
Poland	24	0.00	0.00	0.00	0.42	0.58	0.00
Hungary	25	0.00	0.00	0.00	0.44	0.56	0.00
Spain	26	0.00	0.00	0.00	0.84	0.16	0.00
Latvia	27	0.00	0.00	0.00	0.91	0.09	0.00
United States	28	0.00	0.00	0.00	0.93	0.07	0.00
Russian Federation	29	0.00	0.00	0.00	1.00	0.00	0.00
Portugal	30	0.00	0.00	0.00	1.00	0.00	0.00
Italy	31	0.00	0.00	0.00	1.00	0.00	0.00
Greece	32	0.00	0.00	0.99	0.01	0.00	0.00
Serbia and Montenegro	33	0.00	0.00	1.00	0.00	0.00	0.00
Turkey	34	0.00	0.00	1.00	0.00	0.00	0.00
Uruguay	35	0.00	0.00	1.00	0.00	0.00	0.00
Thailand	36	0.00	0.00	1.00	0.00	0.00	0.00
Mexico	37	0.02	0.98	0.00	0.00	0.00	0.00
Indonesia	38	1.00	0.00	0.00	0.00	0.00	0.00
Tunisia	39	1.00	0.00	0.00	0.00	0.00	0.00
Brazil	40	1.00	0.00	0.00	0.00	0.00	0.00

Probability of masspoint grouping is .95 or more; high degree of confidence

Probability of masspoint grouping is .90 to .95; borderline degree of confidence

Probability of masspoint grouping is less than .90; low degree of confidence

OECD countries are in regular font; partner countries are in *italics* .

Mean Scores on the Four Mathematics Subscales

PISA formed mathematics subscales based on the four major content areas (overarching ideas): Space & Shape, Change & Relationships, Quantity, and Uncertainty. The OECD mean scores on these scales vary slightly from the mean of 500.0 that was set for the overall mathematics scale in 2003. A detailed description of the knowledge and processes associated with the four mathematics subscales may be found in Chapter 1. Appendix A contains a selection of sample items from each subscale, while the descriptions of the proficiency levels in Tables 3.13, 3.14, 3.15 and 3.16 in this chapter provide additional information about each subscale.

Space and Shape Subscale

Questions on the Space & Shape subscale require skills such as recognising and understanding geometric patterns and identifying and manipulating such patterns in abstract and real-life representations. The mean performance of Irish students on the Space & Shape subscale (476.2) is significantly below the OECD mean score of 496.3 (Table 3.3), and is also lower than the Irish mean score for the combined scale.

Table 3.3. Mean Achievement Scores and Standard Deviations on the Space and Shape Subscale – OECD and Partner Countries

	Pop	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*		*Pop*	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*
Hong Kong - Ch	☒	558.4	(4.85)	110.9	(2.90)	▲	Luxembourg	☒	488.2	(1.35)	100.3	(1.23)	▼
Japan	☒	553.2	(4.31)	109.5	(2.87)	▲	*Latvia*	☒	486.4	(4.04)	101.5	(1.71)	▼
Korea	☒	551.7	(3.80)	116.8	(2.49)	▲	Norway	☒	482.7	(2.54)	103.3	(1.26)	▼
Switzerland	☒	539.5	(3.50)	109.8	(2.07)	▲	Hungary	☒	479.1	(3.34)	109.3	(2.22)	▼
Finland	☒	539.0	(2.04)	92.2	(1.25)	▲	Spain	☒	476.5	(2.59)	91.8	(1.37)	▼
Liechtenstein	☐	538.2	(4.58)	106.5	(4.30)	▲	**Ireland**	☒	**476.2**	**(2.43)**	**94.5**	**(1.53)**	▼
Belgium	☒	529.6	(2.26)	110.8	(1.37)	▲	*Russian Fed*	☒	474.3	(4.69)	112.3	(2.00)	▼
Macao - Ch	☐	527.9	(3.29)	97.3	(3.32)	▲	United States	☒	472.0	(2.78)	97.5	(1.35)	▼
Czech Rep	☒	527.4	(4.12)	119.2	(2.32)	▲	Italy	☒	470.3	(3.14)	109.1	(1.76)	▼
Netherlands	☒	526.2	(2.87)	94.0	(2.30)	▲	Portugal	☒	450.2	(3.43)	93.3	(1.69)	▼
New Zealand	☒	524.9	(2.34)	105.7	(1.27)	▲	Greece	☒	437.1	(3.80)	99.8	(1.61)	▼
Australia	☒	520.6	(2.33)	103.6	(1.65)	▲	*Serbia and Monte.*	☒	432.5	(3.91)	95.6	(1.83)	▼
Canada	☒	517.8	(1.81)	95.3	(0.90)	▲	*Thailand*	☐	423.9	(3.35)	90.3	(1.81)	▼
Austria	☒	515.2	(3.48)	111.6	(1.69)	▲	Turkey	×	417.4	(6.35)	101.9	(5.13)	▼
Denmark	☒	512.4	(2.76)	103.3	(1.59)	▲	*Uruguay*	×	412.0	(2.98)	101.1	(1.71)	▼
France	☒	507.6	(2.98)	102.4	(2.01)	▲	Mexico	×	381.7	(3.20)	87.0	(1.45)	▼
Slovak Rep	☒	505.4	(4.01)	116.7	(2.27)	▲	*Indonesia*	×	360.8	(3.70)	88.4	(1.91)	▼
Iceland	☒	503.5	(1.46)	94.3	(1.54)	▲	*Tunisia*	☒	358.9	(2.56)	91.9	(1.71)	▼
Germany	☒	499.6	(3.28)	111.8	(1.87)	o	*Brazil*	×	350.7	(4.11)	96.4	(2.31)	▼
Sweden	☒	498.3	(2.56)	99.9	(1.70)	o	**OECD Total**		**485.8**	**(1.04)**	**112.0**	**(0.69)**	
Poland	☒	490.3	(2.66)	106.9	(1.88)	▼	**OECD Average**		**496.3**	**(0.65)**	**110.1**	**(0.41)**	

☒ >90% of 15-year olds enrolled; ☐ 75-90% of 15-year olds enrolled; × 50-75% of 15-year olds enrolled

■ Mean significantly higher than Ireland; □ Mean not significantly different from Ireland; ▒ Mean significantly lower than Ireland

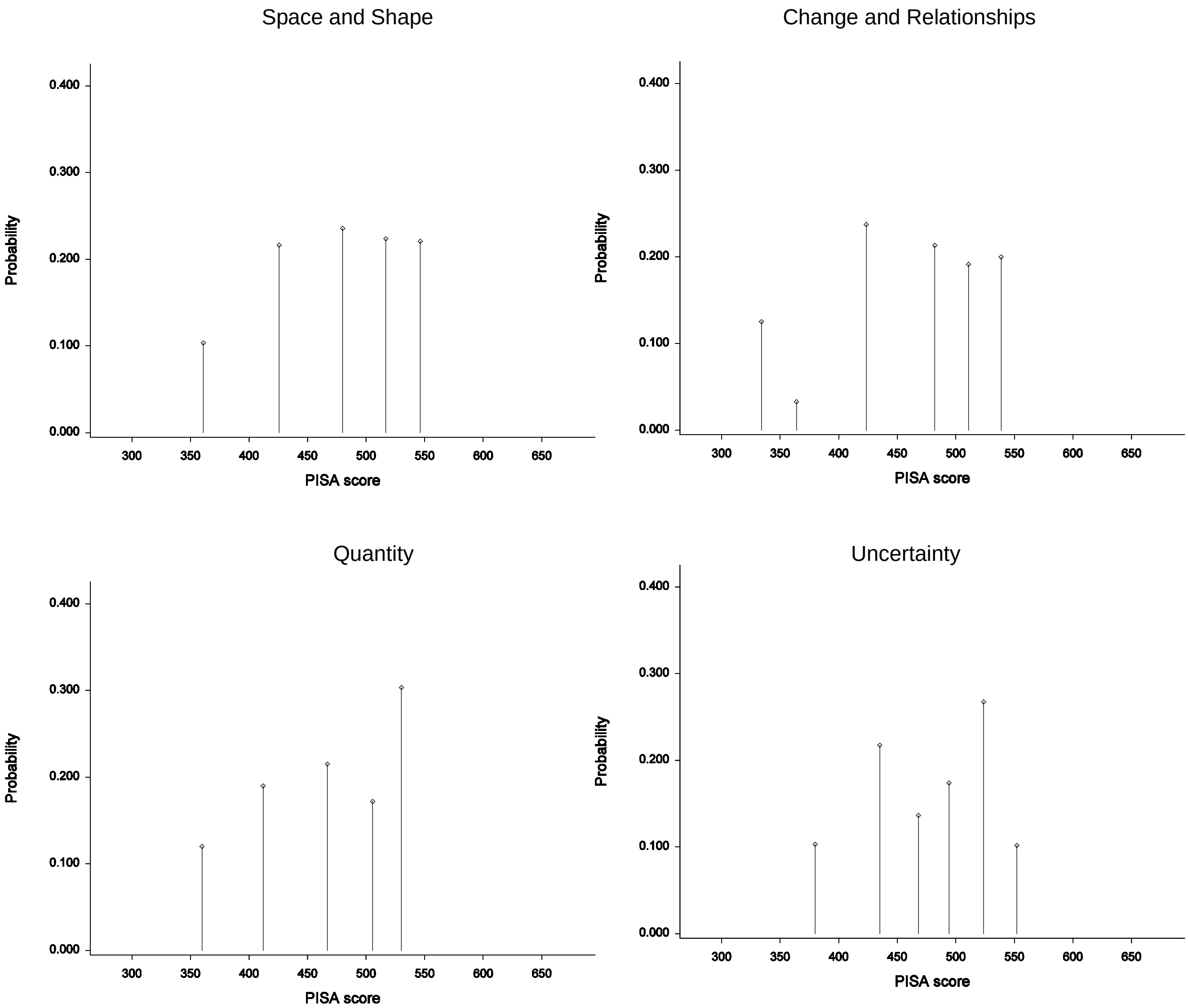
▲ Above OECD average; o At OECD average; ▼ Below OECD average

Note. OECD countries are in regular font; partner countries are in italics. SD = Standard deviation; SE = Standard error.

The column "Pop" is an indicator of the percent of the 15-year-old population enrolled in schools in each country and is based on Column 15 of Table A3.1 in OECD (2004b).

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.

Figure 3.2. *Plots of the NPML Probability Distributions for Scores on Space and Shape (upper left), Change and Relationships (upper right), Quantity (lower left), and Uncertainty (lower right)*



Ireland’s mean score ranks 27th of 40 countries overall (95% confidence interval for Ireland’s ranking = 25th to 29th), and 23rd of 29 OECD countries (95% confidence interval for Ireland’s ranking = 21st to 24th). Ireland’s mean score is significantly lower than that of 22 countries (20 OECD countries), the same as seven countries (five OECD countries), and significantly higher than ten countries (four OECD countries). Ireland scored significantly lower than countries such as Austria, Denmark, and France, which have a score that is not statistically different from Ireland on the combined scale (Table 3.1).

When countries are grouped into distinct performance groupings using the nonparametric maximum likelihood estimation method (see Inset 3.2), five distinct groupings are identified, one fewer than for the combined scale (Table 3.4, and Figure 3.2, upper left). In contrast to the combined scale, more countries have uncertain membership of adjacent groups. Ireland, however, falls with close to 100% confidence into the third performance grouping which has a mean of 480.0. Hence, Ireland’s mean performance does not differ from that of Luxembourg, Norway, Hungary, Spain, the USA, Italy, and partner countries Latvia and the Russian Federation. And, although Austria and Denmark have mean performances that do not differ from Ireland's for the combined scale (Table 3.2), they are in a higher performance grouping (mean of 516.8) for the Space & Shape subscale.

Table 3.4. *Five-point NPML Probability Distribution and Posterior Probabilities for the Space and Shape Subscale – OECD and Partner Countries*

	Rank	*Masspoints*				
PISA score		*361.1*	*425.8*	*480.0*	*516.8*	*546.3*
Proportion		*0.103*	*0.216*	*0.236*	*0.224*	*0.221*
Hong Kong-China	1	0.00	0.00	0.00	0.00	1.00
Japan	2	0.00	0.00	0.00	0.01	0.99
Korea	3	0.00	0.00	0.00	0.01	0.99
Switzerland	4	0.00	0.00	0.00	0.13	0.87
Finland	5	0.00	0.00	0.00	0.14	0.86
Liechtenstein	6	0.00	0.00	0.00	0.17	0.83
Belgium	7	0.00	0.00	0.00	0.62	0.38
Macao-China	8	0.00	0.00	0.00	0.71	0.29
Czech Republic	9	0.00	0.00	0.00	0.73	0.27
Netherlands	10	0.00	0.00	0.00	0.79	0.21
New Zealand	11	0.00	0.00	0.00	0.83	0.17
Australia	12	0.00	0.00	0.00	0.93	0.07
Canada	13	0.00	0.00	0.00	0.96	0.03
Austria	14	0.00	0.00	0.01	0.97	0.02
Denmark	15	0.00	0.00	0.02	0.98	0.01
France	16	0.00	0.00	0.06	0.94	0.00
Slovak Republic	17	0.00	0.00	0.12	0.88	0.00
Iceland	18	0.00	0.00	0.18	0.81	0.00
Germany	19	0.00	0.00	0.43	0.57	0.00
Sweden	20	0.00	0.00	0.52	0.48	0.00
Poland	21	0.00	0.00	0.92	0.08	0.00
Luxembourg	22	0.00	0.00	0.96	0.04	0.00
Latvia	23	0.00	0.00	0.98	0.02	0.00
Norway	24	0.00	0.00	0.99	0.01	0.00
Hungary	25	0.00	0.00	1.00	0.00	0.00
Spain	26	0.00	0.00	1.00	0.00	0.00
Ireland	**27**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**
Russian Federation	28	0.00	0.00	1.00	0.00	0.00
United States	29	0.00	0.00	1.00	0.00	0.00
Italy	30	0.00	0.00	1.00	0.00	0.00
Portugal	31	0.00	0.75	0.25	0.00	0.00
Greece	32	0.00	1.00	0.00	0.00	0.00
Serbia and Montenegro	33	0.00	1.00	0.00	0.00	0.00
Thailand	34	0.00	1.00	0.00	0.00	0.00
Turkey	35	0.00	1.00	0.00	0.00	0.00
Uruguay	36	0.00	1.00	0.00	0.00	0.00
Mexico	37	1.00	0.00	0.00	0.00	0.00
Indonesia	38	1.00	0.00	0.00	0.00	0.00
Tunisia	39	1.00	0.00	0.00	0.00	0.00
Brazil	40	1.00	0.00	0.00	0.00	0.00

Probability of masspoint grouping is .95 or more; high degree of confidence

Probability of masspoint grouping is .90 to .95; borderline degree of confidence

Probability of masspoint grouping is less than .90; low degree of confidence

OECD countries are in regular font; partner countries are in *italics* .

Change and Relationships Subscale

Questions on the Change & Relationships subscale assess understanding of relationships in a variety of forms including symbolic, algebraic, graphical, tabular and geometric, and the ability to translate between these. The mean performance of Irish students on the subscale (506.0) is significantly above the OECD country average score of 498.8. Ireland's mean score ranks 18th of 40 countries overall (95% confidence interval for Ireland's ranking = 15th to 20th), 15th of 29 OECD countries (95% confidence interval of Ireland's ranking = 12th to 17th). It is significantly lower than that of 12 countries (10 OECD countries), the same as that of nine countries (eight OECD countries), and significantly higher than that of 18 countries (10 OECD countries).

Table 3.5. Mean Achievement Scores and Standard Deviations on the Change and Relationships Subscale – OECD and Partner Countries

	Pop	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*
Netherlands	⊠	551.4	(3.12)	94.2	(2.04)	▲
Korea	⊠	547.6	(3.52)	99.5	(2.38)	▲
Finland	⊠	543.1	(2.19)	94.6	(1.36)	▲
Hong Kong - Ch	⊠	539.7	(4.68)	106.0	(2.93)	▲
Liechtenstein	□	539.5	(3.67)	107.1	(3.85)	▲
Canada	⊠	536.7	(1.93)	92.2	(0.94)	▲
Japan	⊠	536.1	(4.33)	112.4	(3.03)	▲
Belgium	⊠	535.3	(2.45)	116.5	(1.64)	▲
New Zealand	⊠	525.7	(2.37)	103.0	(1.46)	▲
Australia	⊠	525.3	(2.30)	97.7	(1.77)	▲
Switzerland	⊠	522.7	(3.65)	111.7	(2.19)	▲
France	⊠	519.7	(2.62)	99.9	(2.09)	▲
Macao - Ch	□	518.8	(3.53)	98.8	(2.87)	▲
Czech Rep	⊠	514.8	(3.50)	100.3	(1.84)	▲
Iceland	⊠	509.5	(1.43)	97.0	(1.18)	▲
Denmark	⊠	509.3	(2.99)	97.7	(1.76)	▲
Germany	⊠	507.2	(3.73)	108.8	(1.71)	▲
Ireland	⊠	**506.0**	**(2.45)**	**87.5**	**(1.40)**	▲
Sweden	⊠	505.1	(2.94)	111.3	(1.91)	▲
Austria	⊠	499.8	(3.60)	102.4	(1.77)	o
Hungary	⊠	494.6	(3.10)	98.8	(2.07)	o
Slovak Rep	⊠	494.4	(3.48)	105.3	(2.34)	o
Norway	⊠	487.7	(2.64)	98.4	(1.30)	▼
Latvia	⊠	487.2	(4.36)	100.8	(1.65)	▼
Luxembourg	⊠	487.0	(1.15)	101.7	(1.01)	▼
United States	⊠	485.5	(3.03)	97.6	(1.61)	▼
Poland	⊠	484.3	(2.70)	99.5	(1.74)	▼
Spain	⊠	480.7	(2.80)	99.1	(1.36)	▼
Russian Fed	⊠	476.8	(4.64)	99.9	(2.13)	▼
Portugal	⊠	467.9	(3.95)	99.3	(2.24)	▼
Italy	⊠	452.1	(3.21)	102.9	(1.89)	▼
Greece	⊠	435.6	(4.31)	107.0	(1.68)	▼
Turkey	×	422.8	(7.57)	120.6	(5.45)	▼
Serbia and Monte.	⊠	419.0	(3.97)	98.5	(1.71)	▼
Uruguay	×	417.0	(3.60)	115.4	(1.70)	▼
Thailand	□	405.0	(3.39)	93.1	(2.12)	▼
Mexico	×	364.1	(4.14)	98.5	(1.93)	▼
Tunisia	⊠	336.6	(2.78)	102.6	(1.87)	▼
Indonesia	×	333.9	(4.58)	105.3	(2.60)	▼
Brazil	×	333.4	(5.99)	124.2	(3.42)	▼
OECD Total		**488.6**	**(1.17)**	**113.0**	**(0.81)**	
OECD Average		**498.8**	**(0.70)**	**109.3**	**(0.46)**	

⊠ >90% of 15-year olds enrolled
□ 75-90% of 15-year olds enrolled
× 50-75% of 15-year olds enrolled

■ (black) Mean significantly higher than Ireland
□ (white) Mean not significantly different from Ireland
(grey) Mean significantly lower than Ireland

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

Note. OECD countries are in regular font; partner countries are in italics. SD = Standard deviation; SE = Standard error.

The column "Pop" is an indicator of the percent of the 15-year-old population enrolled in schools in each country and is based on Column 15 of Table A3.1 in OECD (2004b).

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.

When countries are split into performance groupings using the nonparametric maximum likelihood estimation method (see Inset 3.2), six distinct groups emerge (Table 3.6 and Figure 3.2, upper right). This suggests that the Change & Relationships scale discriminates performance at the country level somewhat better than the Space & Shape subscale. Ireland fares better in this categorisation, falling into the second highest group (mean = 510.8). Its performance on the subscale is indistinguishable from that of the Czech

Republic, Iceland, Denmark, Germany, Sweden (and partner country Macao-China, with 90% probability).

Table 3.6. Six-point NPML Probability Distribution and Posterior Probabilities for the Change and Relationships Subscale – OECD and Partner Countries

	Rank	Masspoints					
PISA score		*334.0*	*364.0*	*423.6*	*482.1*	*510.8*	*538.7*
Proportion		*0.125*	*0.033*	*0.237*	*0.213*	*0.191*	*0.200*
Netherlands	1	0.00	0.00	0.00	0.00	0.00	1.00
Korea	2	0.00	0.00	0.00	0.00	0.00	1.00
Finland	3	0.00	0.00	0.00	0.00	0.00	1.00
Hong Kong-China	4	0.00	0.00	0.00	0.00	0.00	1.00
Liechtenstein	5	0.00	0.00	0.00	0.00	0.00	1.00
Canada	6	0.00	0.00	0.00	0.00	0.01	0.99
Japan	7	0.00	0.00	0.00	0.00	0.01	0.99
Belgium	8	0.00	0.00	0.00	0.00	0.02	0.98
New Zealand	9	0.00	0.00	0.00	0.00	0.40	0.60
Australia	10	0.00	0.00	0.00	0.00	0.43	0.57
Switzerland	11	0.00	0.00	0.00	0.00	0.67	0.33
France	12	0.00	0.00	0.00	0.00	0.86	0.14
Macao-China	13	0.00	0.00	0.00	0.00	0.90	0.10
Czech Republic	14	0.00	0.00	0.00	0.00	0.97	0.03
Iceland	15	0.00	0.00	0.00	0.01	0.99	0.00
Denmark	16	0.00	0.00	0.00	0.01	0.99	0.00
Germany	17	0.00	0.00	0.00	0.02	0.98	0.00
Ireland	**18**	**0.00**	**0.00**	**0.00**	**0.03**	**0.97**	**0.00**
Sweden	19	0.00	0.00	0.00	0.04	0.96	0.00
Austria	20	0.00	0.00	0.00	0.24	0.76	0.00
Hungary	21	0.00	0.00	0.00	0.69	0.31	0.00
Slovak Republic	22	0.00	0.00	0.00	0.70	0.30	0.00
Norway	23	0.00	0.00	0.00	0.97	0.03	0.00
Latvia	24	0.00	0.00	0.00	0.97	0.03	0.00
Luxembourg	25	0.00	0.00	0.00	0.97	0.03	0.00
United States	26	0.00	0.00	0.00	0.99	0.01	0.00
Poland	27	0.00	0.00	0.00	0.99	0.01	0.00
Spain	28	0.00	0.00	0.00	1.00	0.00	0.00
Russian Federation	29	0.00	0.00	0.00	1.00	0.00	0.00
Portugal	30	0.00	0.00	0.00	1.00	0.00	0.00
Italy	31	0.00	0.00	0.65	0.35	0.00	0.00
Greece	32	0.00	0.00	1.00	0.00	0.00	0.00
Turkey	33	0.00	0.00	1.00	0.00	0.00	0.00
Serbia and Montenegro	34	0.00	0.00	1.00	0.00	0.00	0.00
Uruguay	35	0.00	0.00	1.00	0.00	0.00	0.00
Thailand	36	0.00	0.00	1.00	0.00	0.00	0.00
Mexico	37	0.01	0.99	0.00	0.00	0.00	0.00
Tunisia	38	1.00	0.00	0.00	0.00	0.00	0.00
Indonesia	39	1.00	0.00	0.00	0.00	0.00	0.00
Brazil	40	1.00	0.00	0.00	0.00	0.00	0.00

Probability of masspoint grouping is .95 or more; high degree of confidence

Probability of masspoint grouping is .90 to .95; borderline degree of confidence

Probability of masspoint grouping is less than .90; low degree of confidence

OECD countries are in regular font; partner countries are in *italics*.

Quantity Subscale

The Quantity subscale measures concepts relating to the processing and understanding of numbers that are represented in various ways, and assesses understanding of relative size, recognition of numerical patterns, and use of numbers to represent quantities and quantifiable attributes of real-world objects. Ireland's mean score of 501.7 is not significantly different from the OECD average of 500.7 (Table 3.7). Ireland ranks 21st out of 40 countries on the subscale (95% confidence interval of Ireland's ranking = 20th to 23rd), and 18th of 29 OECD countries (95% confidence interval of Ireland's ranking = 17th to 20th). Fifteen countries (12 of the 29 OECD countries) have a score that is significantly higher than that of Ireland; the mean scores of 10 countries (all OECD countries) do not differ from that of Ireland, and the mean scores of 14 countries (six OECD countries) are significantly lower. Thus, Ireland's performance on the Quantity subscale is similar to its performance on the combined mathematics scale.

Table 3.7. Mean Achievement Scores and Standard Deviations on the Quantity Subscale – OECD and Partner Countries

	Pop	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*		*Pop*	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*
Finland	⊠	548.5	(1.83)	83.3	(1.07)	▲	Luxembourg	⊠	501.5	(1.06)	91.1	(1.06)	o
Hong Kong - Ch	⊠	545.2	(4.19)	98.9	(2.62)	▲	Hungary	⊠	496.3	(2.72)	95.1	(1.91)	o
Korea	⊠	537.2	(2.97)	89.9	(1.85)	▲	Norway	⊠	494.2	(2.22)	94.0	(1.11)	▼
Liechtenstein	□	533.5	(4.12)	93.2	(4.51)	▲	Spain	⊠	492.3	(2.53)	96.7	(1.28)	▼
Macao - Ch	□	533.0	(2.99)	87.3	(2.27)	▲	Poland	⊠	491.8	(2.47)	89.0	(1.65)	▼
Switzerland	⊠	532.6	(3.08)	95.6	(1.73)	▲	*Latvia*	⊠	481.7	(3.60)	84.6	(1.41)	▼
Belgium	⊠	529.6	(2.31)	109.5	(1.83)	▲	United States	⊠	476.4	(3.18)	104.9	(1.54)	▼
Netherlands	⊠	528.3	(3.09)	96.7	(2.36)	▲	Italy	⊠	474.8	(3.38)	106.0	(2.01)	▼
Canada	⊠	528.1	(1.85)	93.8	(0.86)	▲	*Russian Fed*	⊠	472.4	(4.04)	92.2	(1.66)	▼
Czech Rep	⊠	528.0	(3.54)	97.8	(2.14)	▲	Portugal	⊠	465.4	(3.51)	93.6	(1.75)	▼
Japan	⊠	526.6	(3.79)	101.7	(2.51)	▲	*Serbia and Monte.*	⊠	456.3	(3.79)	88.7	(1.64)	▼
Australia	⊠	516.9	(2.06)	96.8	(1.53)	▲	Greece	⊠	445.9	(3.97)	100.0	(1.73)	▼
Denmark	⊠	515.6	(2.64)	92.1	(1.60)	▲	*Uruguay*	×	429.7	(3.22)	108.6	(1.60)	▼
Germany	⊠	513.8	(3.37)	105.8	(1.87)	▲	*Thailand*	□	414.8	(3.15)	93.2	(2.11)	▼
Sweden	⊠	513.6	(2.49)	90.1	(1.66)	▲	Turkey	×	413.2	(6.78)	112.0	(5.09)	▼
Iceland	⊠	513.3	(1.50)	96.3	(1.32)	▲	Mexico	×	393.8	(3.94)	95.4	(1.92)	▼
Austria	⊠	513.2	(3.00)	85.8	(1.66)	▲	*Tunisia*	⊠	364.4	(2.79)	88.3	(2.11)	▼
Slovak Rep	⊠	512.5	(3.43)	93.9	(2.32)	▲	*Brazil*	×	359.9	(5.04)	108.9	(3.01)	▼
New Zealand	⊠	511.1	(2.22)	98.8	(1.27)	▲	*Indonesia*	×	357.5	(4.28)	91.4	(2.38)	▼
France	⊠	506.9	(2.49)	95.3	(1.83)	▲	**OECD Total**		**487.3**	**(1.11)**	**107.9**	**(0.73)**	
Ireland	⊠	**501.7**	**(2.48)**	**88.2**	**(1.28)**	o	**OECD Average**		**500.7**	**(0.63)**	**102.3**	**(0.42)**	

⊠	>90% of 15-year olds enrolled	■	Mean significantly higher than Ireland	▲	Above OECD average
□	75-90% of 15-year olds enrolled	□	Mean not significantly different from Ireland	o	At OECD average
×	50-75% of 15-year olds enrolled	▒	Mean significantly lower than Ireland	▼	Below OECD average

Note. OECD countries are in regular font; partner countries are in italics. SD = Standard deviation; SE = Standard error.
The column "Pop" is an indicator of the percent of the 15-year-old population enrolled in schools in each country and is based on Column 15 of Table A3.1 in OECD (2004b).
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.

Following application of the nonparametric maximum likelihood estimation method for grouping countries on the Quantity subscale (Inset 3.2, Table 3.8 and Figure 3.2, lower left), one can distinguish five performance groupings. The distribution of ability groupings is highly negatively skewed; in fact the mean performance of the first 11 countries falls into a single group (with a mean of 530.1). Ireland falls into the second grouping (with a mean of 505.6). Between the first and the second groupings, there are nine countries whose group membership is uncertain.

Table 3.8. *Five-point NPML Probability Distribution and Posterior Probabilities for the Quantity Subscale – OECD and Partner Countries*

	Rank	*Masspoints*				
PISA score		*359.9*	*412.0*	*466.9*	*505.6*	*530.1*
Proportion		*0.120*	*0.190*	*0.215*	*0.172*	*0.303*
Finland	1	0.00	0.00	0.00	0.00	1.00
Hong Kong-China	2	0.00	0.00	0.00	0.00	1.00
Korea	3	0.00	0.00	0.00	0.01	0.99
Liechtenstein	4	0.00	0.00	0.00	0.01	0.99
Macao-China	5	0.00	0.00	0.00	0.02	0.98
Switzerland	6	0.00	0.00	0.00	0.02	0.98
Belgium	7	0.00	0.00	0.00	0.04	0.96
Netherlands	8	0.00	0.00	0.00	0.05	0.95
Canada	9	0.00	0.00	0.00	0.05	0.95
Czech Republic	10	0.00	0.00	0.00	0.05	0.95
Japan	11	0.00	0.00	0.00	0.07	0.93
Australia	12	0.00	0.00	0.00	0.41	0.59
Denmark	13	0.00	0.00	0.00	0.49	0.51
Germany	14	0.00	0.00	0.00	0.61	0.39
Sweden	15	0.00	0.00	0.00	0.59	0.41
Iceland	16	0.00	0.00	0.00	0.62	0.38
Austria	17	0.00	0.00	0.00	0.63	0.37
Slovak Republic	18	0.00	0.00	0.00	0.66	0.34
New Zealand	19	0.00	0.00	0.00	0.73	0.27
France	20	0.00	0.00	0.00	0.88	0.12
Ireland	**21**	**0.00**	**0.00**	**0.00**	**0.96**	**0.04**
Luxembourg	22	0.00	0.00	0.00	0.96	0.04
Hungary	23	0.00	0.00	0.03	0.96	0.01
Norway	24	0.00	0.00	0.06	0.93	0.01
Spain	25	0.00	0.00	0.12	0.88	0.00
Poland	26	0.00	0.00	0.14	0.85	0.00
Latvia	27	0.00	0.00	0.87	0.13	0.00
United States	28	0.00	0.00	0.98	0.02	0.00
Italy	29	0.00	0.00	0.99	0.01	0.00
Russian Federation	30	0.00	0.00	0.99	0.01	0.00
Portugal	31	0.00	0.00	1.00	0.00	0.00
Serbia and Montenegro	32	0.00	0.00	1.00	0.00	0.00
Greece	33	0.00	0.03	0.97	0.00	0.00
Uruguay	34	0.00	0.99	0.01	0.00	0.00
Thailand	35	0.00	1.00	0.00	0.00	0.00
Turkey	36	0.00	1.00	0.00	0.00	0.00
Mexico	37	0.01	0.99	0.00	0.00	0.00
Tunisia	38	1.00	0.00	0.00	0.00	0.00
Brazil	39	1.00	0.00	0.00	0.00	0.00
Indonesia	40	1.00	0.00	0.00	0.00	0.00

Probability of masspoint grouping is .95 or more; high degree of confidence

Probability of masspoint grouping is .90 to .95; borderline degree of confidence

Probability of masspoint grouping is less than .90; low degree of confidence

OECD countries are in regular font; partner countries are in *italics* .

Uncertainty Subscale

The Uncertainty subscale measures concepts relating to data and chance, and is therefore associated with the study of statistics and probability. Ireland's mean score of 517.2 is significantly above the OECD average of 502.0 by 15.2 score points or about one-sixth of a standard deviation (Table 3.9). Ireland ranks 13th of 40 countries on this subscale (95% confidence interval for Ireland's ranking = 12th to 16th), and 10th of 29 OECD countries (95% confidence interval for Ireland's ranking = 10th to 13th). Nine countries (seven OECD countries) have a mean score that is significantly higher than Ireland's, eight countries (seven OECD countries) have mean scores that do not differ from Ireland's, and 22 countries (14 OECD countries) have mean scores that are significantly lower. In contrast to the relatively poor Irish performance on the Space & Shape subscale, Ireland's performance on the Uncertainty subscale is on a par with countries that have combined mathematics scores that are above the OECD average, such as Japan, Belgium, and Switzerland.

Table 3.9. Mean Achievement Scores and Standard Deviations on the Uncertainty Subscale – OECD and Partner Countries

	Pop	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*		*Pop*	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*
Hong Kong - Ch	⊠	558.3	(4.56)	100.7	(3.05)	▲	Germany	⊠	492.5	(3.29)	97.6	(1.66)	▼
Netherlands	⊠	549.3	(2.99)	89.6	(2.02)	▲	Luxembourg	⊠	492.1	(1.06)	95.6	(0.97)	▼
Finland	⊠	544.8	(2.09)	84.5	(1.13)	▲	United States	⊠	491.5	(2.97)	98.5	(1.49)	▼
Canada	⊠	541.6	(1.83)	87.2	(0.94)	▲	Hungary	⊠	489.0	(2.63)	85.6	(1.78)	▼
Korea	⊠	538.3	(3.03)	88.9	(1.95)	▲	Spain	⊠	489.0	(2.42)	87.7	(1.40)	▼
New Zealand	⊠	532.2	(2.30)	99.2	(1.32)	▲	Slovak Republic	⊠	475.8	(3.21)	87.0	(1.84)	▼
Macao - Ch	□	531.6	(3.21)	87.6	(2.61)	▲	*Latvia*	⊠	473.8	(3.28)	83.9	(1.42)	▼
Australia	⊠	530.9	(2.21)	97.5	(1.63)	▲	Portugal	⊠	470.6	(3.41)	82.9	(1.78)	▼
Japan	⊠	527.9	(3.88)	98.3	(2.58)	▲	Italy	⊠	462.6	(3.03)	95.2	(1.74)	▼
Iceland	⊠	527.8	(1.50)	95.0	(1.44)	▲	Greece	⊠	458.4	(3.53)	88.2	(1.51)	▼
Belgium	⊠	525.7	(2.21)	106.0	(1.47)	▲	Turkey	×	442.6	(6.21)	98.0	(5.01)	▼
Liechtenstein	□	523.4	(3.68)	96.2	(3.67)	▲	*Russian Fed*	⊠	436.5	(4.02)	89.6	(1.60)	▼
Ireland	⊠	**517.2**	**(2.65)**	**88.8**	**(1.36)**	▲	*Serbia and Monte.*	⊠	427.9	(3.49)	83.2	(1.49)	▼
Switzerland	⊠	516.5	(3.28)	100.2	(2.09)	▲	*Thailand*	□	422.7	(2.53)	72.9	(1.81)	▼
Denmark	⊠	515.6	(2.78)	91.5	(1.60)	▲	*Uruguay*	×	418.6	(3.11)	98.0	(1.71)	▼
Norway	⊠	512.8	(2.59)	98.0	(1.15)	▲	Mexico	×	389.8	(3.26)	80.2	(1.55)	▼
Sweden	⊠	510.8	(2.72)	101.2	(1.69)	▲	*Indonesia*	×	384.5	(2.86)	66.1	(1.55)	▼
France	⊠	506.1	(2.39)	91.7	(1.75)	o	*Brazil*	×	376.6	(3.93)	83.6	(2.68)	▼
Czech Rep	⊠	500.3	(3.11)	91.1	(1.70)	o	*Tunisia*	⊠	363.3	(2.30)	70.5	(1.71)	▼
Austria	⊠	493.8	(3.13)	94.5	(1.71)	▼	**OECD Total**		**491.8**	**(1.07)**	**101.6**	**(0.66)**	
Poland	⊠	493.5	(2.35)	85.1	(1.68)	▼	**OECD Average**		**502.0**	**(0.61)**	**98.6**	**(0.39)**	

⊠	>90% of 15-year olds enrolled	■	Mean significantly higher than Ireland	▲	Above OECD average
□	75-90% of 15-year olds enrolled	□	Mean not significantly different from Ireland	o	At OECD average
×	50-75% of 15-year olds enrolled	▒	Mean significantly lower than Ireland	▼	Below OECD average

Note. OECD countries are in regular font; partner countries are in italics. SD = Standard deviation; SE = Standard error.

The column "Pop" is an indicator of the percent of the 15-year-old population enrolled in schools in each country and is based on Column 15 of Table A3.1 in OECD (2004b).

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.

Table 3.10. Six-point NPML Probability Distribution and Posterior Probabilities for the Uncertainty Subscale – OECD and Partner Countries

	Rank	*Masspoints*					
PISA score		*380.0*	*435.1*	*468.1*	*494.2*	*523.8*	*552.1*
Proportion		*0.103*	*0.217*	*0.136*	*0.174*	*0.267*	*0.102*
Hong Kong-China	1	0.00	0.00	0.00	0.00	0.00	1.00
Netherlands	2	0.00	0.00	0.00	0.00	0.02	0.98
Finland	3	0.00	0.00	0.00	0.00	0.12	0.88
Canada	4	0.00	0.00	0.00	0.00	0.36	0.64
Korea	5	0.00	0.00	0.00	0.00	0.69	0.31
New Zealand	6	0.00	0.00	0.00	0.00	0.97	0.03
Macao-China	7	0.00	0.00	0.00	0.00	0.98	0.02
Australia	8	0.00	0.00	0.00	0.00	0.98	0.02
Japan	9	0.00	0.00	0.00	0.00	0.99	0.00
Iceland	10	0.00	0.00	0.00	0.00	0.99	0.00
Belgium	11	0.00	0.00	0.00	0.00	1.00	0.00
Liechtenstein	12	0.00	0.00	0.00	0.00	1.00	0.00
Ireland	**13**	**0.00**	**0.00**	**0.00**	**0.02**	**0.98**	**0.00**
Switzerland	14	0.00	0.00	0.00	0.02	0.98	0.00
Denmark	15	0.00	0.00	0.00	0.03	0.97	0.00
Norway	16	0.00	0.00	0.00	0.10	0.90	0.00
Sweden	17	0.00	0.00	0.00	0.22	0.78	0.00
France	18	0.00	0.00	0.00	0.70	0.30	0.00
Czech Republic	19	0.00	0.00	0.00	0.97	0.03	0.00
Austria	20	0.00	0.00	0.01	0.99	0.00	0.00
Poland	21	0.00	0.00	0.01	0.99	0.00	0.00
Germany	22	0.00	0.00	0.01	0.99	0.00	0.00
Luxembourg	23	0.00	0.00	0.01	0.99	0.00	0.00
United States	24	0.00	0.00	0.01	0.99	0.00	0.00
Hungary	25	0.00	0.00	0.03	0.97	0.00	0.00
Spain	26	0.00	0.00	0.03	0.97	0.00	0.00
Slovak Republic	27	0.00	0.00	0.86	0.14	0.00	0.00
Latvia	28	0.00	0.00	0.93	0.07	0.00	0.00
Portugal	29	0.00	0.00	0.98	0.02	0.00	0.00
Italy	30	0.00	0.01	0.99	0.00	0.00	0.00
Greece	31	0.00	0.05	0.95	0.00	0.00	0.00
Turkey	32	0.00	0.99	0.01	0.00	0.00	0.00
Russian Federation	33	0.00	1.00	0.00	0.00	0.00	0.00
Serbia and Montenegro	34	0.00	1.00	0.00	0.00	0.00	0.00
Thailand	35	0.00	1.00	0.00	0.00	0.00	0.00
Uruguay	36	0.00	1.00	0.00	0.00	0.00	0.00
Mexico	37	1.00	0.00	0.00	0.00	0.00	0.00
Indonesia	38	1.00	0.00	0.00	0.00	0.00	0.00
Brazil	39	1.00	0.00	0.00	0.00	0.00	0.00
Tunisia	40	1.00	0.00	0.00	0.00	0.00	0.00

Probability of masspoint grouping is .95 or more; high degree of confidence

Probability of masspoint grouping is .90 to .95; borderline degree of confidence

Probability of masspoint grouping is less than .90; low degree of confidence

OECD countries are in regular font; partner countries are in *italics* .

Looking at the performance of OECD and partner countries with Ireland as a reference group and using the same estimation method as before (Inset 3.2), six distinct performance groupings emerge (Table 3.10 and Figure 3.2, lower right), suggesting that the subscale discriminates performance at the country level to a reasonable degree. Just two countries, Hong Kong-China, and the Netherlands, fall with close to 100% confidence into the top group (mean of 552.1); the three countries below the Netherlands cannot be classified with a high degree of certainty into either the first or the second performance grouping. Ireland falls with 98% confidence into the second performance grouping (mean = 523.8), and does not differ from New Zealand, Australia, Japan, Iceland, Belgium, Switzerland, Denmark, and Norway, and partner countries Macao-China and Liechtenstein.

Inset 3.3. Interpreting Proficiency Levels in PISA 2003

What are PISA proficiency levels? The application of techniques associated with Item Response Theory to the PISA achievement data means that it is possible to generate a criterion-referenced interpretation of student performance. Item response techniques enable test item difficulty and student performance to be placed on the same scale. The development of the proficiency levels involved establishing appropriate cut-off points, and describing the skills and knowledge demonstrated by students at each proficiency level. The process of developing proficiency levels was an iterative one in which members of the PISA mathematics and technical expert teams worked together to identify and describe the skills associated with each level. Because proficiency levels for reading literacy were established in 2000, it was possible to apply these to the achievement estimates in the same domain for 2003. Proficiency levels for mathematics were established for the first time using the 2003 data.

PISA proficiency levels for both reading and mathematics were defined in such a way that a student at the bottom of a level has an average probability of .50 of succeeding on the items at that level. Application of this criterion, and a proviso that proficiency levels should be of fixed width, led to the establishment of a response probability convention of .62 (that is, the probability of a student at a particular point on the scale responding correctly to an item at that point is .62). The resulting cut-off points are given in Tables 3.11 (Mathematics) and Table 3.22 (Reading). The term ‘below Level 1’ refers to students who did not meet the criterion for Level 1 (i.e., the estimated probability of these students responding correctly to items at the bottom of Level 1 is less than .50). PISA does not measure what students below Level 1 can accomplish, although it is acknowledged that such students may have mathematics or reading skills other than those assessed. The OECD (2004b) has suggested that students scoring at or below Level 1 in mathematics may not have the mathematical skills they need in adult life. PISA does not describe the upper limits reached by students at the top levels on the scales (i.e., students at Level 6 in mathematics, and Level 5 in reading may have additional higher-level skills not assessed by PISA).

How should PISA proficiency levels be interpreted? PISA proficiency levels should be interpreted with reference to the knowledge and skills associated with the items at each proficiency level (see Tables 3.11, 3.13, 3.14, 3.15, 3.16, and Figure 3.3 for mathematics, and Tables 3.21 and 3.22 and Figure 3.5 for reading literacy). However, levels can be interpreted in statistical terms along the following lines:

- All students within a level are expected to respond correctly to at least half of the items at that level (since the average probability of succeeding on an item is set at .50 for students at the bottom of the level);
- Students at the bottom of a level have a .62 chance of correctly answering the easiest items on that level and a .42 chance of answering the hardest items;
- Students at the top of a level have a .62 chance of correctly answering the most difficult items at that level, and a .78 chance of answering the easiest items; and
- Students just below the top of a level are expected to respond correctly to less than 50% on the items at the next highest level.

Why are there six levels for mathematics but only five for reading? The proficiency levels for mathematics and reading cannot be directly compared. Given the manner in which the test items were distributed along scales in the two domains and the skills required at various points on the scales, mathematics proficiency is better described in six clusters or levels (with a width of about 62 score points), and reading in five (with a width of about 72 score points).

Performance on the Mathematics Proficiency Levels

To represent degrees of proficiency along the combined mathematics scale and the four mathematics subscales, each was divided into six levels (see Inset 3.3 and Table 3.11).

Table 3.11. Descriptions of Proficiency Levels on the Combined Mathematics Scale, and Percentages of Students Achieving at Each Level – Ireland and OECD

		*Ireland **		*OECD***	
Level	*Brief Description – Students at this level are likely to be able to:*	*%*	*(SE)*	*%*	*(SE)*
Level 6 (above 668.7)	Conceptualise, generalise, and utilise information based on investigations and modelling of complex problem situations; link different information sources and representations and flexibly translate among them; demonstrate advanced mathematical thinking and reasoning, and apply this insight along with a mastery of symbolic and formal mathematical operations and relationships to develop new approaches and strategies for attacking novel situations; formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments, and the appropriateness of these to the original situations.	2.2	(0.33)	4.0	(0.10)
Level 5 (606.6 to 668.7)	Develop and work with models for complex situations, identifying constraints and specifying assumptions; select, compare, and evaluate appropriate problem-solving strategies for dealing with complex problems; work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insight pertaining to these situations; and reflect on their actions and formulate and communicate their interpretations and reasoning.	9.1	(0.76)	10.6	(0.13)
Level 4 (544.4 to 606.6)	Work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions; select and integrate different representations, including symbolic ones, linking them directly to aspects of real-world situations; utilise well-developed skills and reason flexibly, with some insight, in these contexts; and construct and communicate explanations based on their own interpretations, arguments, and actions.	20.2	(1.06)	19.1	(0.17)
Level 3 (482.4 to 544.4)	Execute clearly described procedures, including those that require sequential decisions; select and apply simple problem-solving strategies; interpret and use representations based on different information sources and reason directly from them and develop short communications reporting their interpretations, results and reasoning.	28.0	(0.82)	23.7	(0.18)
Level 2 (420.4 to 482.4)	Interpret and recognise situations in contexts that require no more than direct inference; extract relevant information from a single source and make use of a single representational mode; employ basic algorithms, formulae, procedures, or conventions, and demonstrate direct reasoning and make literal interpretations of the results.	23.6	(0.83)	21.1	(0.15)
Level 1 (358.3 to 420.4)	Complete tasks involving familiar contexts where all relevant information is present and the questions are clearly defined; identify information and carry out routine procedures according to direct instructions in explicit situations; and perform actions that are obvious and follow immediately from the given stimuli.	12.1	(0.84)	13.2	(0.16)
Below Level 1 (less than 358.3)	Has a less than .50 chance of responding correctly to Level 1 tasks. Mathematics skills not assessed by PISA.	4.7	(0.57)	8.2	(0.17)
	Total	100.0		100.0	

*N (Ireland) = 3880. **Denotes OECD average percent.

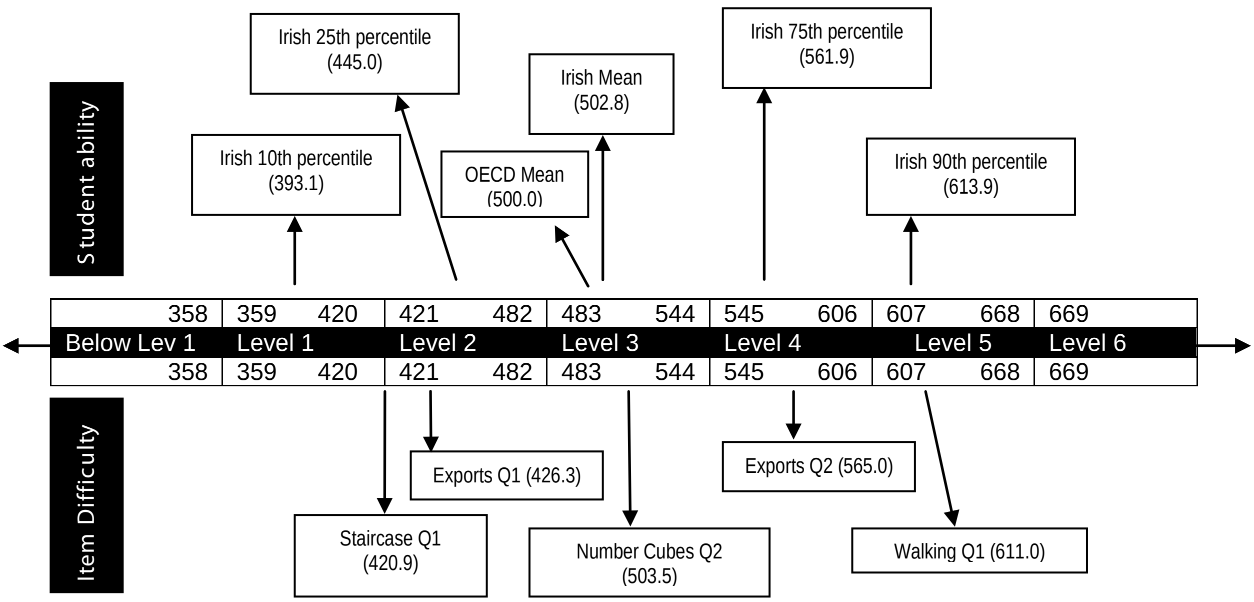
Note. Students at a level have at least a 50% chance of correctly answering all items at that level.

On the combined mathematics scale, students who achieve at Level 6, the highest level, are likely to complete at least 50% of the most complex PISA mathematics tasks successfully. Such tasks require mastery of a range of symbolic and formal mathematical operations and relationships in order to develop new approaches and strategies for novel situations. Students at this level can also formulate and precisely communicate their actions and reflections regarding their findings, interpretations, and arguments. In contrast, students who achieve at Level 1, the lowest level, are likely to succeed only on the more basic mathematics tasks, in contexts where all relevant information is present and the questions are clearly defined. They can perform actions that are obvious and follow immediately from the given stimuli. Between these two extremes, the likely proficiency of students progresses from basic to moderate to proficient. A more complete description of the PISA proficiency levels for the combined mathematics scale may be found in Table 3.11.

Some students were unable to demonstrate proficiency on Level 1 tasks (i.e., their pattern of response indicated that they would not be expected to successfully solve half of the tasks drawn from Level 1). These students fall into the category 'below Level 1'. They may lack the basic mathematical knowledge and skills to progress through the education system, and are likely to be at risk in the transition from education to work, and in their educational and occupational outcomes.

Figure 3.1 shows the relationship between Irish student proficiency and item difficulty for mathematics through a selection of the sample items shown in Appendix A and a selection of key performance benchmarks. For example, the second item from the unit 'Number Cubes' has an item difficulty (503.5) around the Irish student mean (502.8). The second item from the unit 'Exports' has an item difficulty (565.0) that is close to the score of Irish students at the 75th percentile (561.9).

Figure 3.3. The PISA 2003 Combined Mathematics Scale: Cut-points for Proficiency Levels, Scores of Students in Ireland at Key Markers, and Difficulties of Selected Items



Countries with the highest mean scores on combined mathematics generally have the highest percentages of students scoring at proficiency Levels 5 and 6 (Table 3.12). For example, close to 31% of students in Hong Kong-China have scores that are at Level 5 or 6 (over twice the OECD average of 14.6%), while about a quarter of students in Finland, Korea and the Netherlands are attaining Levels 5 or 6. In contrast, in the lowest achieving countries (Mexico, Indonesia, Tunisia, Brazil), less than 1% of students attain Levels 5 or 6. While less than 11% of students in Hong Kong-China, Finland, Korea and the Netherlands score at or below Level 1 (half the OECD average), over three-quarters of students in Indonesia, Tunisia and Brazil score at these levels.

Table 3.12. Percentage of Students at Each Proficiency Level on the Combined Mathematics Scale – OECD and Partner Countries

	< Level 1 (below 358.3)		*Level 1 (358.3 to 420.4)*		*Level 2 (420.4 to 482.4)*		*Level 3 (482.4 to 544.4)*		*Level 4 (544.4 to 606.6)*		*Level 5 (606.6 to 668.7)*		*Level 6 (above 668.7)*	
	%	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*
Hong Kong - China	3.9	(0.72)	6.5	(0.64)	13.9	(1.00)	20.0	(1.25)	25.0	(1.17)	20.2	(1.00)	10.5	(0.94)
Finland	1.5	(0.23)	5.3	(0.38)	16.0	(0.57)	27.7	(0.65)	26.1	(0.89)	16.7	(0.64)	6.7	(0.46)
Korea	2.5	(0.32)	7.1	(0.65)	16.6	(0.80)	24.1	(0.98)	25.0	(1.08)	16.7	(0.81)	8.1	(0.93)
Netherlands	2.6	(0.65)	8.4	(0.95)	18.0	(1.11)	23.0	(1.14)	22.6	(1.34)	18.2	(1.09)	7.3	(0.58)
Liechtenstein	4.8	(1.33)	7.5	(1.66)	17.3	(2.78)	21.6	(2.54)	23.2	(3.09)	18.3	(3.22)	7.3	(1.73)
Japan	4.7	(0.65)	8.6	(0.72)	16.3	(0.80)	22.4	(1.02)	23.6	(1.24)	16.1	(0.96)	8.2	(1.14)
Canada	2.4	(0.26)	7.7	(0.36)	18.3	(0.61)	26.2	(0.67)	25.1	(0.60)	14.8	(0.55)	5.5	(0.45)
Belgium	7.2	(0.56)	9.3	(0.49)	15.9	(0.65)	20.1	(0.71)	21.0	(0.62)	17.5	(0.69)	9.0	(0.48)
Macao - China	2.3	(0.60)	8.8	(1.34)	19.6	(1.40)	26.8	(1.77)	23.7	(1.71)	13.8	(1.55)	4.8	(0.96)
Switzerland	4.9	(0.45)	9.6	(0.57)	17.5	(0.80)	24.3	(0.98)	22.5	(0.72)	14.2	(1.05)	7.0	(0.90)
Australia	4.3	(0.45)	10.0	(0.51)	18.6	(0.62)	24.0	(0.71)	23.3	(0.64)	14.0	(0.53)	5.8	(0.45)
New Zealand	4.9	(0.44)	10.1	(0.63)	19.2	(0.71)	23.2	(0.90)	21.9	(0.80)	14.1	(0.60)	6.6	(0.44)
Czech Republic	5.0	(0.69)	11.6	(0.90)	20.1	(0.96)	24.3	(0.95)	20.8	(0.87)	12.9	(0.80)	5.3	(0.53)
Iceland	4.5	(0.40)	10.5	(0.55)	20.2	(1.02)	26.1	(0.88)	23.2	(0.81)	11.7	(0.61)	3.7	(0.36)
Denmark	4.7	(0.50)	10.7	(0.62)	20.6	(0.89)	26.2	(0.88)	21.9	(0.83)	11.8	(0.86)	4.1	(0.50)
France	5.6	(0.68)	11.0	(0.77)	20.2	(0.82)	25.9	(0.99)	22.1	(0.97)	11.6	(0.72)	3.5	(0.40)
Sweden	5.6	(0.52)	11.7	(0.60)	21.7	(0.84)	25.5	(0.95)	19.8	(0.81)	11.6	(0.57)	4.1	(0.49)
Austria	5.6	(0.70)	13.2	(0.84)	21.6	(0.90)	24.9	(1.14)	20.5	(0.84)	10.5	(0.85)	3.7	(0.52)
Germany	9.2	(0.84)	12.4	(0.81)	19.0	(1.05)	22.6	(0.82)	20.6	(1.02)	12.2	(0.87)	4.1	(0.48)
Ireland	**4.7**	**(0.57)**	**12.1**	**(0.84)**	**23.6**	**(0.83)**	**28.0**	**(0.82)**	**20.2**	**(1.06)**	**9.1**	**(0.76)**	**2.2**	**(0.33)**
Slovak Republic	6.7	(0.85)	13.2	(0.86)	23.5	(0.88)	24.9	(1.08)	18.9	(0.82)	9.8	(0.68)	2.9	(0.38)
Norway	6.9	(0.50)	13.9	(0.82)	23.7	(1.16)	25.2	(1.01)	18.9	(1.00)	8.7	(0.57)	2.7	(0.35)
Luxembourg	7.4	(0.41)	14.3	(0.65)	22.9	(0.87)	25.9	(0.79)	18.7	(0.85)	8.5	(0.59)	2.4	(0.31)
Poland	6.8	(0.61)	15.2	(0.76)	24.8	(0.75)	25.3	(0.94)	17.7	(0.89)	7.8	(0.49)	2.3	(0.31)
Hungary	7.8	(0.80)	15.2	(0.81)	23.8	(1.05)	24.3	(0.93)	18.2	(0.90)	8.2	(0.73)	2.5	(0.42)
Spain	8.1	(0.66)	14.9	(0.87)	24.7	(0.78)	26.7	(1.02)	17.7	(0.65)	6.5	(0.62)	1.4	(0.25)
Latvia	7.6	(0.86)	16.1	(1.08)	25.5	(1.17)	26.3	(1.15)	16.6	(1.17)	6.3	(0.70)	1.6	(0.36)
United States	10.2	(0.80)	15.5	(0.81)	23.9	(0.80)	23.8	(0.79)	16.6	(0.73)	8.0	(0.53)	2.0	(0.36)
Russian Fed.	11.4	(1.03)	18.8	(1.09)	26.4	(1.13)	23.1	(1.02)	13.2	(0.92)	5.4	(0.58)	1.6	(0.38)
Portugal	11.3	(1.11)	18.8	(0.99)	27.1	(0.99)	24.0	(1.03)	13.4	(0.94)	4.6	(0.47)	0.8	(0.16)
Italy	13.2	(1.19)	18.7	(0.93)	24.7	(1.03)	22.9	(0.84)	13.4	(0.73)	5.5	(0.43)	1.5	(0.19)
Greece	17.8	(1.21)	21.2	(1.15)	26.3	(1.04)	20.2	(1.01)	10.6	(0.87)	3.4	(0.53)	0.6	(0.17)
Serbia and Mont.	17.6	(1.35)	24.5	(1.08)	28.6	(1.16)	18.9	(1.11)	8.1	(0.88)	2.1	(0.41)	0.2	(0.10)
Turkey	27.7	(2.01)	24.6	(1.33)	22.1	(1.12)	13.5	(1.27)	6.8	(1.05)	3.1	(0.82)	2.4	(1.02)
Uruguay	26.3	(1.30)	21.8	(0.80)	24.2	(0.89)	16.8	(0.68)	8.2	(0.65)	2.3	(0.33)	0.5	(0.17)
Thailand	23.8	(1.28)	30.2	(1.25)	25.4	(1.12)	13.7	(0.85)	5.3	(0.53)	1.5	(0.31)	0.2	(0.10)
Mexico	38.1	(1.71)	27.9	(1.02)	20.8	(0.87)	10.1	(0.84)	2.7	(0.39)	0.4	(0.10)	0.02	(0.01)
Indonesia	50.5	(2.08)	27.6	(1.05)	14.8	(1.07)	5.5	(0.71)	1.4	(0.39)	0.2	(0.09)	0.01	n/a
Tunisia	51.1	(1.37)	26.9	(0.95)	14.7	(0.75)	5.7	(0.61)	1.4	(0.30)	0.2	(0.12)	0.01	n/a
Brazil	53.3	(1.94)	21.9	(1.09)	14.1	(0.86)	6.8	(0.78)	2.7	(0.47)	0.9	(0.36)	0.28	(0.2)
OECD Total	**11.0**	**(0.32)**	**14.6**	**(0.32)**	**21.2**	**(0.28)**	**22.4**	**(0.32)**	**17.6**	**(0.25)**	**9.6**	**(0.19)**	**3.5**	**(0.19)**
OECD Average	**8.2**	**(0.17)**	**13.2**	**(0.16)**	**21.1**	**(0.15)**	**23.7**	**(0.18)**	**19.1**	**(0.17)**	**10.6**	**(0.13)**	**4.0**	**(0.10)**

Note. OECD countries are in regular font, partner countries are in *italics* . Countries are ordered in descending order of mean score on the combined mathematics scale.

An examination of the distribution of Irish students across proficiency levels on the combined mathematics scale shows that the percentage of students who are at or below Level 1 is below the OECD average (16.8% compared to 21.4%). However, fewer Irish students attain Levels 5 and 6 either (11.3% compared to 14.6%). Thus, Ireland's moderate performance may be attributed to the comparatively low performance of high achievers rather than to the low performance of low achievers.

When proficiency levels are applied to each of the mathematics subscales, notable differences in student performance become apparent. Table 3.13 shows the percentages of Irish students at each proficiency level of the Space & Shape subscale, compared to the OECD average percentages, with a description of task characteristics associated with each level (see also Appendix A for sample items from this subscale).

Table 3.13. Descriptions of Proficiency Levels on the Space and Shape Subscale, and Percentages of Students Achieving at Each Level – Ireland and OECD

		*Ireland **		*OECD***	
Level	*Brief Description – Students at this level are likely to be able to:*	*%*	*(SE)*	*%*	*(SE)*
Level 6 (above 668.7)	Solve complex problems involving multiple representations and often involving sequential calculation processes; identify and extract relevant information and link different but related information; use reasoning, significant insight and reflection; generalise results and findings, communicate solutions and provide explanations and argumentation.	1.8	(0.24)	5.8	(0.10)
Level 5 (606.6 to 668.7)	Solve problems that require appropriate assumptions to be made, or that involve working with assumptions provided; use well-developed spatial reasoning, argument and insight to identify relevant information and to interpret and link different representations; work strategically and carry out multiple and sequential processes.	6.8	(0.64)	10.4	(0.11)
Level 4 (544.4 to 606.6)	Solve problems that involve visual and spatial reasoning and argumentation in unfamiliar contexts; link and integrate different representations; carry out sequential processes; apply well-developed skills in spatial visualisation and interpretation.	15.4	(0.77)	17.2	(0.15)
Level 3 (482.4 to 544.4)	Solve problems that involve elementary visual and spatial reasoning in familiar contexts; link different representations of familiar objects; use elementary problem-solving skills (devising simple strategies); apply simple algorithms.	23	(1.02)	21.5	(0.16)
Level 2 (420.4 to 482.4)	Solve problems involving a single mathematical representation where the mathematical content is direct and clearly presented; use basic mathematical thinking and conventions in familiar contexts.	25.4	(0.87)	20.4	(0.14)
Level 1 (358.3 to 420.4)	Solve simple problems in a familiar context using familiar pictures or drawings of geometric objects and applying counting or basic calculation skills.	16.9	(1.15)	14.2	(0.16)
Below Level 1 (less than 358.3)	Has a less than .50 chance of responding correctly to Level 1 tasks. Mathematics skills not assessed by PISA.	10.7	(0.78)	10.6	(0.19)
	Total	100.0		100.0	

*N (Ireland) = 3880. **Denotes OECD average percent.

Note. Students at a level have at least a 50% chance of correctly answering all items at that level.

The overall poor performance of Irish students in this domain is echoed in the distribution of student scores across proficiency levels. For example, 27.6% of Irish students (compared with an OECD average of 24.8%) score at or below Level 1, and over half (53.0%) score at or below Level 2 (compared to an OECD average of 45.2%). At the upper end of the scale, less than 9% of Irish students reach Levels 5 or 6, compared to almost twice this percentage (16.2%) on average for OECD countries. Comparing the percentages for Ireland for the combined scale and the Space & Shape subscale, the difference is marked at the lower end where about 40% are at or below Level 2 on the combined scale compared to 53.0% on the Space & Shape subscale.

The picture for Ireland on the Change & Relationships subscale is similar to that for the combined scale. Table 3.14 shows the percentages of Irish students at each proficiency level of the subscale, compared to the OECD average percentages, with a description of task characteristics at each level (see also Appendix A for sample items from the Change & Relationships subscale). About 16% of Irish students (compared with an OECD average of 23.2%) are at or below Level 1 and about 39% at or below Level 2 (compared with an OECD average of 43.0%). At the upper end of the scale, 12.5% of Irish students (compared with an OECD average of 16.4%) are at Level 5 or 6, once again indicating the comparatively weaker performance of higher achievers in Ireland.

Table 3.14. Descriptions of Proficiency Levels on the Change and Relationships Subscale, and Percentages of Students Achieving at Each Level – Ireland and OECD

		*Ireland **		*OECD***	
Level	*Brief Description – Students at this level are likely to be able to:*	*%*	*(SE)*	*%*	*(SE)*
Level 6 (above 668.7)	Use significant insight, abstract reasoning and argumentation skills and technical knowledge and conventions to solve problems and to generalise mathematical solutions to complex real-world problems.	2.3	(0.35)	5.3	(0.11)
Level 5 (606.6 to 668.7)	Solve problems by making advanced use of algebraic and other formal mathematical expressions and models; link formal mathematical representations to complex real-world situations; use complex and multi-step problem-solving skills, reflect on and communica	10.2	(0.63)	11.1	(0.13)
Level 4 (544.4 to 606.6)	Understand and work with multiple representations, including explicitly mathematical models of real-world situations to solve practical problems; employ considerable flexibility in interpretation and reasoning, including in unfamiliar contexts, and commun	21.6	(0.85)	18.5	(0.20)
Level 3 (482.4 to 544.4)	Solve problems that involve working with multiple related representations (a text, a graph, a table, a formula), including some interpretation, reasoning in familiar contexts, and communication of argument.	27	(1.07)	22.0	(0.19)
Level 2 (420.4 to 482.4)	Work with simple algorithms, formulae and procedures to solve problems; link text with a single representation (graph, table, simple formula); use interpretation and reasoning skills at an elementary level.	22.6	(0.84)	19.8	(0.14)
Level 1 (358.3 to 420.4)	Locate relevant information in a simple table or graph; follow direct and simple instructions to read information directly from a simple table or graph in a standard or familiar form; perform simple calculations involving relationships between two familia	11.2	(0.86)	13.0	(0.15)
Below Level 1 (less than 358.3)	Has a less than .50 chance of responding correctly to Level 1 tasks. Mathematics skills not assessed by PISA.	5.1	(0.51)	10.2	(0.19)
	Total	100.0		100.0	

*N (Ireland) = 3880. **Denotes OECD average percent.

Note. Students at a level have at least a 50% chance of correctly answering all items at that level.

Table 3.15 shows the percentage of Irish students and the OECD average percentage of students attaining each proficiency level of the Quantity subscale, together with a description of task characteristics (see also Appendix A). The distribution of performance is again characterised by a slightly lower percentage of Irish students at, or below, Level 1, compared with the OECD average (17.9% compared to 21.3%). On the other hand, the percentage of Irish students at Levels 5 and 6 (11.7%) is also below the OECD average (15.0%). This is similar to the percentage of students at Levels 5 and 6 on the combined scale (11.3%).

Table 3.15. Descriptions of Proficiency Levels on the Quantity Subscale, and Percentages of Students Achieving at Each Level – Ireland and OECD

		*Ireland **		*OECD***	
Level	*Brief Description – Students at this level are likely to be able to:*	*%*	*(SE)*	*%*	*(SE)*
Level 6 (above 668.7)	Conceptualise and work with models of complex mathematical processes and relationships; work with formal and symbolic expressions; use advanced reasoning skills to devise strategies for solving problems and to link multiple contexts; use sequential calculation process; formulate conclusions, arguments and precise explanations.	2.2	(0.36)	4.0	(0.09)
Level 5 (606.6 to 668.7)	Work effectively with models of more complex situations to solve problems; use well-developed reasoning skills, insight and interpretation with different representations; carry out sequential processes; communicate reasoning and argument.	9.5	(0.62)	11.0	(0.11)
Level 4 (544.4 to 606.6)	Work effectively with simple models of complex situations; use reasoning skills in a variety of contexts; interpret different representations of the same situation; analyse and apply quantitative relationships; use a variety of calculation skills to solve problems.	20.6	(0.84)	19.9	(0.17)
Level 3 (482.4 to 544.4)	Use simple problem-solving strategies including reasoning in familiar contexts; interpret tables to locate information; carry out explicitly described calculations including sequential processes.	26.9	(1.06)	23.7	(0.21)
Level 2 (420.4 to 482.4)	Interpret simple tables to identify and extract relevant information; carry out basic arithmetic calculations; interpret and work with simple quantitative relationships.	23.0	(1.00)	20.1	(0.15)
Level 1 (358.3 to 420.4)	Solve problems of the most basic type in which all relevant information is explicitly presented, the situation is straightforward and very limited in scope, the required computational activity is obvious and the mathematical task is basic, such as a simple arithmetic operation.	12.3	(0.85)	12.5	(0.15)
Below Level 1 (less than 358.3)	Has a less than .50 chance of responding correctly to Level 1 tasks. Mathematics skills not assessed by PISA.	5.6	(0.57)	8.8	(0.18)
	Total	100.0		100.0	

* N (Ireland) = 3880. **Denotes OECD average percent

Note. Students at a level have at least a 50% chance of correctly answering all items at that level.

Table 3.16 shows the percentages of Irish students and OECD average percentages for each proficiency level for the Uncertainty subscale together with a description of task characteristics (see Appendix A for sample items). Consistent with Ireland's high average performance on this subscale, proportionately more Irish students (16.4%) are attaining Levels 5 and 6 (compared to an OECD average of 14.8%). In fact, twice as many Irish students attain Level 5 or 6 on the Uncertainty subscale compared to the Space & Shape subscale. Similarly, 13.8% of Irish students (compared to an OECD average of 20.7%) are at or below Level 1, and about 35% score at or below Level 2 (a considerable improvement on the 53.0% observed for the Space & Shape subscale).

Table 3.16. Descriptions of Proficiency Levels on the Uncertainty Subscale, and Percentages of Students Achieving at Each Level – Ireland and OECD

		*Ireland **		*OECD***	
Level	*Brief Description – Students at this level are likely to be able to:*	*%*	*(SE)*	*%*	*(SE)*
Level 6 (above 668.7)	Use high-level thinking and reasoning skills in statistical or probabilistic contexts to create mathematical representations of real-world situations; use insight and reflection to solve problems, and to formulate and communicate arguments and explanations.	4	(0.39)	4.2	(0.10)
Level 5 (606.6 to 668.7)	Apply probabilistic and statistical knowledge in problem situations that are somewhat structured and where the mathematical representation is partially apparent; use reasoning and insight to interpret and analyse given information, to develop appropriate models and to perform sequential calculation processes; communicate reasons and arguments.	12.4	(0.72)	10.6	(0.14)
Level 4 (544.4 to 606.6)	Use basic statistical and probabilistic concepts combined with numerical reasoning in less familiar contexts to solve simple problems; carry out multi-step or sequential calculation processes; use and communicate argumentation based on interpretation of data.	22.0	(0.93)	19.2	(0.17)
Level 3 (482.4 to 544.4)	Interpret statistical information and data, and link different information sources; engage basic reasoning with simple probability concepts, symbols and conventions and communication of reasoning.	26.5	(0.93)	23.8	(0.17)
Level 2 (420.4 to 482.4)	Locate statistical information presented in familiar graphical form; understand basic statistical concepts and conventions.	21.2	(0.86)	21.5	(0.16)
Level 1 (358.3 to 420.4)	Understand and use basic probabilistic ideas in familiar experimental contexts.	10.2	(0.74)	13.3	(0.17)
Below Level 1 (less than 358.3)	Has a less than .50 chance of responding correctly to Level 1 tasks. Mathematics skills not assessed by PISA.	3.6	(0.45)	7.4	(0.15)
	Total	100.0		100.0	

* N (Ireland) = 3880. **Denotes OECD average percent

Note. Students at a level have at least a 50% chance of correclty answering all items at that level.

Variation in Performance on the Combined Mathematics Scale

In the previous section, the performance of students on the mathematics proficiency levels was described in terms of specified levels of knowledge and skills in the context of *absolute* benchmarks (the proficiency levels). In this section, the focus shifts to a consideration of the *relative* dispersion of scores in Ireland and in other countries, including the gap between the best and poorest performing students. Such a gap may be interpreted as an indicator of similarity in educational outcomes, with a small gap indicating higher levels of equality in outcomes, and a large gap indicating inequality. As can be seen in Table 3.1, the standard deviation for the combined mathematics scale for Ireland (85.3) is comparatively small, indicating a relatively small dispersion in achievement. The standard deviation in Germany, a country with a combined mathematics score not significantly different from that of Ireland, is 102.6, suggesting a greater spread in achievement. The focus of this section is on performance at the 10th and 90th percentiles, that is, the performance of students near the top and bottom of the achievement distribution. However, Table 3.17 also presents data for the 5th, 25th, 75th, and 95th percentiles.

Table 3.17. *Scores at the 5th, 10th, 25th, 75th, 90th, and 95th Percentiles on the Combined Mathematics Scale – OECD and Partner Countries*

	5th		*10th*		*25th*		*75th*		*90th*		*95th*	
	Score	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*
Hong Kong - China	373.8	(11.05)	417.0	(8.02)	484.8	(6.91)	621.8	(3.74)	671.8	(4.10)	699.5	(3.97)
Finland	406.4	(3.83)	438.0	(2.77)	488.2	(2.21)	602.6	(2.32)	651.7	(2.83)	680.2	(3.13)
Korea	387.8	(4.61)	422.8	(4.46)	479.3	(3.74)	606.1	(4.22)	659.2	(5.37)	690.2	(6.83)
Netherlands	385.2	(6.86)	415.4	(5.84)	470.9	(5.44)	608.3	(3.84)	656.5	(3.21)	683.5	(3.43)
Liechtenstein	361.9	(19.68)	408.0	(9.77)	469.9	(7.58)	608.6	(7.91)	655.3	(9.53)	686.4	(16.38)
Japan	360.9	(8.24)	401.7	(6.26)	467.2	(5.37)	605.1	(4.36)	659.6	(6.14)	690.2	(6.58)
Canada	386.2	(3.05)	419.3	(2.54)	473.9	(2.19)	593.3	(2.13)	644.2	(2.58)	672.7	(3.39)
Belgium	333.8	(6.53)	380.7	(4.61)	456.2	(3.44)	611.2	(2.49)	664.4	(2.35)	693.4	(2.38)
Macao - China	382.3	(8.76)	414.5	(5.97)	467.2	(4.41)	587.3	(4.01)	639.1	(5.48)	668.4	(8.28)
Switzerland	358.7	(4.80)	395.7	(4.16)	460.8	(3.57)	595.0	(4.89)	652.1	(5.23)	684.0	(6.84)
Australia	364.3	(4.44)	398.6	(3.43)	459.8	(2.75)	591.6	(2.50)	644.7	(3.04)	675.7	(3.53)
New Zealand	358.5	(4.07)	394.3	(3.89)	455.2	(2.91)	593.0	(2.21)	650.0	(3.20)	682.3	(2.91)
Czech Republic	358.0	(6.25)	391.7	(5.72)	449.4	(4.55)	584.4	(3.98)	641.0	(4.35)	671.9	(4.89)
Iceland	362.4	(4.05)	396.1	(2.74)	454.2	(2.81)	578.4	(1.95)	629.2	(3.02)	657.9	(3.77)
Denmark	360.7	(4.39)	395.8	(4.53)	453.2	(3.67)	578.2	(3.14)	631.5	(3.65)	662.0	(4.73)
France	352.4	(5.96)	388.7	(5.56)	449.1	(3.74)	575.2	(3.01)	627.7	(3.58)	656.2	(3.46)
Sweden	352.7	(5.29)	387.1	(4.38)	446.1	(3.02)	575.6	(3.19)	630.5	(3.82)	661.9	(4.80)
Austria	353.4	(6.64)	384.4	(4.44)	439.4	(4.02)	571.4	(4.18)	626.2	(3.96)	658.2	(4.96)
Germany	324.0	(6.08)	363.0	(5.65)	432.2	(4.66)	578.3	(3.48)	632.3	(3.50)	661.7	(3.64)
Ireland	**360.4**	**(4.68)**	**393.1**	**(3.21)**	**445.0**	**(3.38)**	**561.9**	**(3.01)**	**613.9**	**(3.59)**	**641.0**	**(3.30)**
Slovak Republic	342.4	(6.91)	378.5	(5.81)	435.6	(4.57)	564.6	(3.78)	619.1	(3.49)	648.4	(4.07)
Norway	343.5	(3.96)	375.9	(3.42)	432.9	(2.87)	560.0	(3.32)	613.6	(3.56)	644.7	(3.92)
Luxembourg	338.5	(3.87)	372.7	(2.69)	430.2	(2.15)	557.2	(1.91)	611.4	(3.20)	641.4	(2.72)
Poland	343.4	(5.78)	376.0	(3.62)	428.2	(3.13)	552.8	(2.87)	607.4	(3.34)	639.9	(3.50)
Hungary	335.3	(5.62)	369.6	(4.23)	426.1	(3.01)	555.9	(3.90)	610.7	(4.71)	643.8	(4.59)
Spain	335.0	(5.13)	368.6	(3.54)	426.2	(2.98)	546.4	(3.12)	597.4	(3.50)	626.0	(3.70)
Latvia	339.2	(5.90)	370.8	(5.14)	423.5	(3.90)	543.5	(4.72)	596.4	(4.43)	626.3	(4.97)
United States	323.0	(4.88)	356.5	(4.55)	418.0	(3.69)	549.7	(3.36)	607.4	(3.87)	638.0	(5.14)
Russian Federation	318.5	(5.46)	350.8	(4.96)	405.8	(4.83)	530.1	(4.95)	588.1	(5.28)	622.4	(6.10)
Portugal	320.9	(6.26)	351.9	(5.25)	406.0	(4.96)	526.1	(3.52)	579.9	(3.29)	609.9	(3.72)
Italy	307.2	(6.39)	342.4	(5.86)	400.5	(4.34)	530.2	(3.01)	589.1	(3.63)	623.2	(3.74)
Greece	287.6	(5.39)	323.5	(5.13)	382.4	(4.57)	507.9	(4.28)	565.9	(5.25)	597.8	(5.10)
Serbia and Montenegro	298.9	(4.37)	328.6	(4.47)	378.6	(3.96)	493.1	(4.78)	546.4	(5.05)	579.2	(5.29)
Turkey	269.7	(5.76)	300.2	(5.01)	350.8	(5.26)	484.9	(8.53)	559.7	(14.23)	613.6	(22.75)
Uruguay	255.3	(4.30)	291.3	(3.80)	353.3	(4.07)	490.7	(3.77)	550.0	(4.36)	583.4	(4.67)
Thailand	289.9	(3.95)	316.3	(3.10)	360.5	(2.92)	469.3	(3.75)	526.0	(4.70)	560.0	(6.43)
Mexico	247.1	(5.39)	276.1	(4.70)	326.6	(4.32)	443.6	(4.46)	497.1	(4.69)	526.9	(5.65)
Indonesia	233.2	(5.22)	260.5	(4.81)	306.0	(3.49)	411.5	(4.77)	465.8	(6.50)	498.8	(7.69)
Tunisia	228.6	(3.80)	256.3	(3.51)	303.0	(2.55)	411.6	(3.59)	465.8	(4.78)	501.4	(6.80)
Brazil	202.6	(5.98)	233.3	(5.32)	285.8	(4.56)	419.3	(6.18)	487.8	(9.53)	528.3	(11.35)
OECD Total	**314.6**	**(2.14)**	**351.9**	**(1.73)**	**418.0**	**(1.61)**	**562.6**	**(1.10)**	**622.3**	**(1.27)**	**655.2**	**(1.80)**
OECD Average	**331.7**	**(1.30)**	**369.0**	**(1.11)**	**432.4**	**(0.93)**	**570.5**	**(0.70)**	**628.3**	**(0.74)**	**660.2**	**(0.95)**

Note. OECD countries are in regular font, partner countries are in *italics* . Countries are ordered in descending order of mean score on the combined mathematics scale.

The OECD average difference between the 10th and 90th percentiles is 259.3 score points. The difference for Ireland is smaller, at 220.9 score points, again indicating a relatively narrow spread in achievement. The OECD average difference is larger than the difference between the highest and lowest country means (194.4 points), demonstrating that differences amongst students within countries are larger than differences between countries. Ireland again stands in contrast to Germany, a country with a similar mean score but a large difference between these two points (269.3 score points). The differences between the 10th and 90th percentiles for Finland and Korea are also comparatively small, indicating that high achievement and homogeneity in achievement outcomes are not incompatible outcomes of education systems.

Variation in Performance on the Mathematics Subscales

Table 3.18 shows the scores of Irish students at the 5th, 10th, 25th, 75th and 90th percentiles, and the OECD averages at these percentiles, for each of the four mathematics subscales. The overall weak performance in Space & Shape is again indicated by the fact that at the 90th percentile, the Irish score (599.0) is almost 40 score points lower than the OECD average (638.6); at the 10th percentile, the score for Irish students (354.0) and the OECD average (354.4) are about the same. In contrast, the score for Ireland at the 10th percentile for the Change & Relationships subscale (393.2) is 37.1 points higher than the OECD average at that marker (356.1), but the score at the 90th percentile (617.5) is almost 20 points lower than the corresponding OECD average (637.4). The Irish score on the 90th percentile for the Quantity scale (614.9) is about 14 points lower than the OECD average (629.1); at the 10th percentile, the score is about 22 points higher. Ireland compares most favourably on the Uncertainty subscale, where the score at the 10th percentile (402.5) is 28.4 points higher than the OECD average (374.1), and there is little difference between Ireland's score (632.5) and the OECD average score (629.1) at the 90th percentile.

Table 3.18. Scores at the 5th, 10th, 25th, 75th, 90th, and 95th Percentiles on the Four Mathematics Subscales – Ireland and OECD

	Space and Shape				*Change and Relationships*			
	Ireland		*OECD*		*Ireland*		*OECD*	
Percentile	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*
5th	323.6	(4.42)	315.5	(1.43)	357.2	(4.43)	313.1	1.50
10th	354.0	(3.63)	354.4	(1.19)	393.2	(4.56)	356.1	1.22
25th	412.3	(3.32)	420.9	(0.87)	448.2	(3.38)	426.4	0.95
75th	541.9	(2.87)	572.3	(0.68)	567.5	(2.77)	575.5	0.70
90th	599.0	(4.54)	638.6	(0.83)	617.5	(2.64)	637.4	0.79
95th	631.9	(4.16)	676.9	(1.02)	645.1	(3.56)	672.4	0.93
	Quantity				*Uncertainty*			
	Ireland		*OECD*		*Ireland*		*OECD*	
Percentile	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*
5th	352.8	(5.29)	325.2	(1.39)	370.8	(5.58)	338.8	(1.14)
10th	387.6	(4.34)	365.7	(1.17)	402.5	(4.52)	374.1	(1.00)
25th	442.0	(3.42)	433.3	(0.87)	455.9	(3.48)	434.1	(0.85)
75th	563.9	(2.97)	573.0	(0.60)	579.8	(3.44)	571.2	(0.66)
90th	614.9	(3.14)	629.1	(0.68)	632.5	(3.42)	629.1	(0.69)
95th	644.0	(3.16)	660.9	(0.80)	661.4	(3.55)	662.1	(0.93)

ACHIEVEMENT IN READING LITERACY

Twenty-eight of the test items used in the reading literacy test in 2000 were administered again in 2003. The items were selected with reference to the range of item difficulties and aspects of the framework, so that there are similar proportions of items falling into the three reading process areas (Retrieve, Interpret, Reflect) in 2000 and 2003. It is possible in 2003 to re-apply the cut-points associated with the reading proficiency levels that were used in 2000 (see Inset 3.3) and to compare performances in the two years. Although PISA 2000 reported performance on a combined reading literacy scale and on three subscales, results for 2003 are reported on a combined scale only. In this section, performance in 2003 is described, while in the last section in the chapter, performance in 2003 is compared with performance in 2000.

Mean Scores on the Reading Literacy Scale

Ireland achieved a mean score of 515.5 on the reading literacy scale (Table 3.19) This is significantly higher than the OECD country mean of 494.2. Ireland's ranking is 7th of 40 countries (95% confidence interval for Ireland's ranking = 6th to 10th), and 6th out of 29 OECD countries (95% confidence interval for Ireland's ranking = 6th to 8th). Just three countries have mean scores that are significantly higher than that of Ireland (Finland, Korea, Canada).

Seven countries have mean scores that are not significantly different, while 29 countries have mean scores that are significantly lower.

As with mathematics, country means are more clustered at the upper than at the lower end of the distribution. For example, just 28.0 score points separate the top seven performing countries, while 66.4 score points separate the seven lowest performing countries. The low standard deviation associated with the mean score for Ireland (86.5 compared to an OECD average of 100.2) indicates a relatively narrow dispersion of achievement. This stands in contrast to countries such as New Zealand and Belgium, which have mean scores that do not differ from Ireland's, but have standard deviations which are larger.

Finland, Korea and Canada emerge as consistently high performers in both reading literacy and combined mathematics (Tables 3.1, 3.19), while the performance of some countries (e.g., Japan) is relatively poorer in reading literacy than in combined mathematics.

It is noteworthy that six of the seven countries with mean scores that are not significantly different from Ireland's in reading literacy (Australia, New Zealand, the Netherlands, Belgium, Liechtenstein, and Hong Kong-China) all have significantly higher mean scores in combined mathematics.

Table 3.19. Mean Achievement Scores and Standard Deviations on the Reading Literacy Scale – OECD and Partner Countries

	Pop	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*		*Pop*	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*
Finland	☒	543.5	(1.64)	81.0	(1.13)	▲	Austria	☒	490.7	(3.76)	103.1	(2.26)	o
Korea	☒	534.1	(3.09)	82.6	(2.03)	▲	*Latvia*	☒	490.6	(3.67)	90.4	(1.75)	o
Canada	☒	527.9	(1.75)	88.5	(0.93)	▲	Czech Rep	☒	488.5	(3.46)	95.5	(2.39)	o
Australia	☒	525.4	(2.13)	97.4	(1.52)	▲	Hungary	☒	481.9	(2.47)	92.0	(1.82)	▼
Liechtenstein	☐	525.1	(3.58)	89.8	(3.35)	▲	Spain	☒	480.5	(2.60)	95.4	(1.48)	▼
New Zealand	☒	521.6	(2.46)	104.6	(1.46)	▲	Luxembourg	☒	479.4	(1.48)	99.7	(1.03)	▼
Ireland	☒	**515.5**	**(2.63)**	**86.5**	**(1.75)**	▲	Portugal	☒	477.6	(3.73)	92.7	(2.13)	▼
Sweden	☒	514.3	(2.42)	95.6	(1.91)	▲	Italy	☒	475.7	(3.04)	100.7	(2.16)	▼
Netherlands	☒	513.1	(2.85)	84.8	(2.05)	▲	Greece	☒	472.3	(4.10)	104.5	(1.95)	▼
Hong Kong - Ch	☒	509.5	(3.69)	84.8	(2.74)	▲	Slovak Rep	☒	469.2	(3.12)	92.5	(2.03)	▼
Belgium	☒	507.0	(2.58)	110.0	(2.14)	▲	*Russian Fed*	☒	442.2	(3.94)	93.3	(1.76)	▼
Norway	☒	499.7	(2.78)	102.5	(1.81)	o	Turkey	×	441.0	(5.79)	95.3	(4.11)	▼
Switzerland	☒	499.1	(3.28)	94.8	(1.90)	o	*Uruguay*	×	434.1	(3.43)	121.5	(2.01)	▼
Japan	☒	498.1	(3.92)	105.5	(2.53)	o	Thailand	☐	419.9	(2.81)	78.1	(1.50)	▼
Macao - Ch	☐	497.6	(2.16)	66.9	(1.86)	o	*Serbia and Monte.*	☒	411.7	(3.56)	81.5	(1.63)	▼
Poland	☒	496.6	(2.88)	95.9	(1.76)	o	*Brazil*	×	402.8	(4.58)	111.3	(2.30)	▼
France	☒	496.2	(2.68)	97.0	(2.17)	o	Mexico	×	399.7	(4.09)	95.1	(1.93)	▼
United States	☒	495.2	(3.22)	101.2	(1.44)	o	*Indonesia*	×	381.6	(3.38)	76.3	(1.79)	▼
Denmark	☒	492.3	(2.82)	88.3	(1.75)	o	*Tunisia*	☒	374.6	(2.81)	95.7	(1.78)	▼
Iceland	☒	491.7	(1.56)	98.3	(1.37)	o	**OECD Total**		**487.7**	**(1.18)**	**103.8**	**(0.66)**	
Germany	☒	491.4	(3.39)	109.1	(2.25)	o	**OECD Average**		**494.2**	**(0.64)**	**100.2**	**(0.39)**	

☒ >90% of 15-year olds enrolled; ☐ 75-90% of 15-year olds enrolled; × 50-75% of 15-year olds enrolled

Black: Mean significantly higher than Ireland; White: Mean not significantly different from Ireland; Grey: Mean significantly lower than Ireland

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

Note. OECD countries are in regular font; partner countries are in italics. SD = Standard deviation; SE = Standard error.

The column "Pop" is an indicator of the percent of the 15-year-old population enrolled in schools in each country and is based on Column 15 of Table A3.1 in OECD (2004b).

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.

When country mean scores are compared using the nonparametric maximum likelihood estimation method, which clusters countries into distinct performance groupings (see Inset 3.3), 10 distinct groupings emerge (Table 3.20 and Figure 3.4). This stands in contrast to the six distinct groupings found for the combined mathematics scale (Table 3.2), and suggests that the PISA assessment of reading discriminates better between countries than PISA combined mathematics. Ireland falls into the third highest of these groupings (which has a mean of 511.5) and is indistinguishable from Sweden, the Netherlands, Belgium, and partner country Hong Kong-China. Korea and Finland are in the top group (mean = 536.1). Countries in the bottom three groupings are Thailand, Serbia and Montenegro, Brazil, Mexico, Indonesia, and Tunisia.

Performance on the Reading Proficiency Levels

To represent degrees of proficiency along the reading literacy scale each was divided into five levels using the same cut-points that were developed in 2000 (see Inset 3.3 and Table 3.21).

Table 3.20. *Ten-point NPML Probability Distribution and Posterior Probabilities for the Reading Literacy Scale: OECD and Partner Countries*

	Rank	*Masspoints*									
PISA score		*378.7*	*401.5*	*414.9*	*440.0*	*475.4*	*491.3*	*497.1*	*511.5*	*524.7*	*536.1*
Proportion		*0.047*	*0.093*	*0.050*	*0.149*	*0.157*	*0.159*	*0.150*	*0.100*	*0.065*	*0.030*
Finland	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Korea	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.98
Canada	3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.98	0.02
Australia	4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
Liechtenstein	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
New Zealand	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.99	0.00
Ireland	**7**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.99**	**0.01**	**0.00**
Sweden	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
Netherlands	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
Hong Kong-China	10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
Belgium	11	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.98	0.00	0.00
Norway	12	0.00	0.00	0.00	0.00	0.00	0.03	0.97	0.00	0.00	0.00
Switzerland	13	0.00	0.00	0.00	0.00	0.00	0.04	0.96	0.00	0.00	0.00
Japan	14	0.00	0.00	0.00	0.00	0.00	0.08	0.92	0.00	0.00	0.00
Macao-China	15	0.00	0.00	0.00	0.00	0.00	0.11	0.89	0.00	0.00	0.00
Poland	16	0.00	0.00	0.00	0.00	0.00	0.19	0.81	0.00	0.00	0.00
France	17	0.00	0.00	0.00	0.00	0.00	0.23	0.77	0.00	0.00	0.00
United States	18	0.00	0.00	0.00	0.00	0.00	0.37	0.63	0.00	0.00	0.00
Denmark	19	0.00	0.00	0.00	0.00	0.00	0.78	0.22	0.00	0.00	0.00
Iceland	20	0.00	0.00	0.00	0.00	0.00	0.84	0.16	0.00	0.00	0.00
Germany	21	0.00	0.00	0.00	0.00	0.00	0.87	0.13	0.00	0.00	0.00
Austria	22	0.00	0.00	0.00	0.00	0.00	0.91	0.09	0.00	0.00	0.00
Latvia	23	0.00	0.00	0.00	0.00	0.00	0.92	0.08	0.00	0.00	0.00
Czech Republic	24	0.00	0.00	0.00	0.00	0.00	0.98	0.02	0.00	0.00	0.00
Hungary	25	0.00	0.00	0.00	0.00	0.93	0.07	0.00	0.00	0.00	0.00
Spain	26	0.00	0.00	0.00	0.00	0.99	0.01	0.00	0.00	0.00	0.00
Luxembourg	27	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Portugal	28	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Italy	29	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Greece	30	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Slovak Republic	31	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Russian Federation	32	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
Turkey	33	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
Uruguay	34	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
Thailand	35	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Serbia and Mont.	36	0.00	0.01	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Brazil	37	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mexico	38	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Indonesia	39	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tunisia	40	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

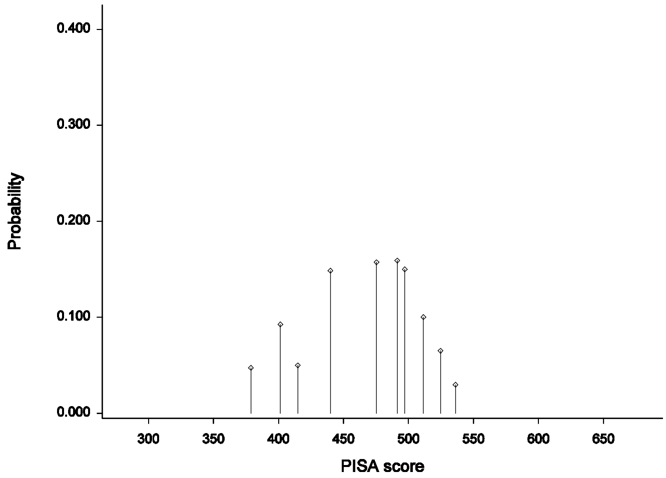
Probability of masspoint grouping is .95 or more; high degree of confidence

Probability of masspoint grouping is .90 to .95; borderline degree of confidence

Probability of masspoint grouping is less than .90; low degree of confidence

OECD countries are in regular font; partner countries are in *italics* .

Figure 3.4. Plot of the NPML Probability Distributions for Reading Country Scores



On the reading literacy scale, students who achieve at the highest level (Level 5), are capable of completing the most complex reading tasks in PISA, working with dense text, and making high-level inferences. The scale is unbounded at the upper level, so that some students may be capable of even more complex reading tasks than those assessed in PISA. In contrast, students who achieve at the lowest level (Level 1), are most likely to succeed only on the more basic reading tasks, such as locating a single piece of information in an elementary level text. Levels 1 to 5, therefore, represent a gradual increase in the reading skills of students, from basic, to moderately complex, to complex, to very complex. A more complete description of the PISA proficiency levels for the combined reading literacy scale is presented in Table 3.21.

As in combined mathematics, some students were unable to demonstrate proficiency on the easiest PISA reading tasks (i.e., their pattern of responses indicated that they would not be expected to successfully solve half of the tasks drawn from Level 1). These students fall into the category 'below Level 1'. Such students are likely to have serious difficulties in applying what reading skills they have to advance and extend their knowledge, and hence may be at a disadvantage in advancing their educational and occupational careers.

Figure 3.5 shows the relationship between Irish student proficiency and item difficulty for reading using a selection of the sample items released following PISA 2000 (Cosgrove, Sofroniou, Kelly, & Shiel, 2003) and a selection of key performance benchmarks. For example, the second item from the unit 'Labour' has an item difficulty (631.1) above the Irish 90th percentile (622.1). The first item from the unit 'Labour' and the second item from the unit 'Gift' have item difficulties (476.6 and 447.5, respectively) that fall either side of the Irish 25th percentile (460.2).

Proportionately fewer students in Ireland (11.0%) scored at or below Level 1 on the reading literacy scale, compared to the OECD average (19.1%), and slightly more scored at Level 4 or 5 (35.5% compared to the OECD average of 29.6%, see Table 3.22).

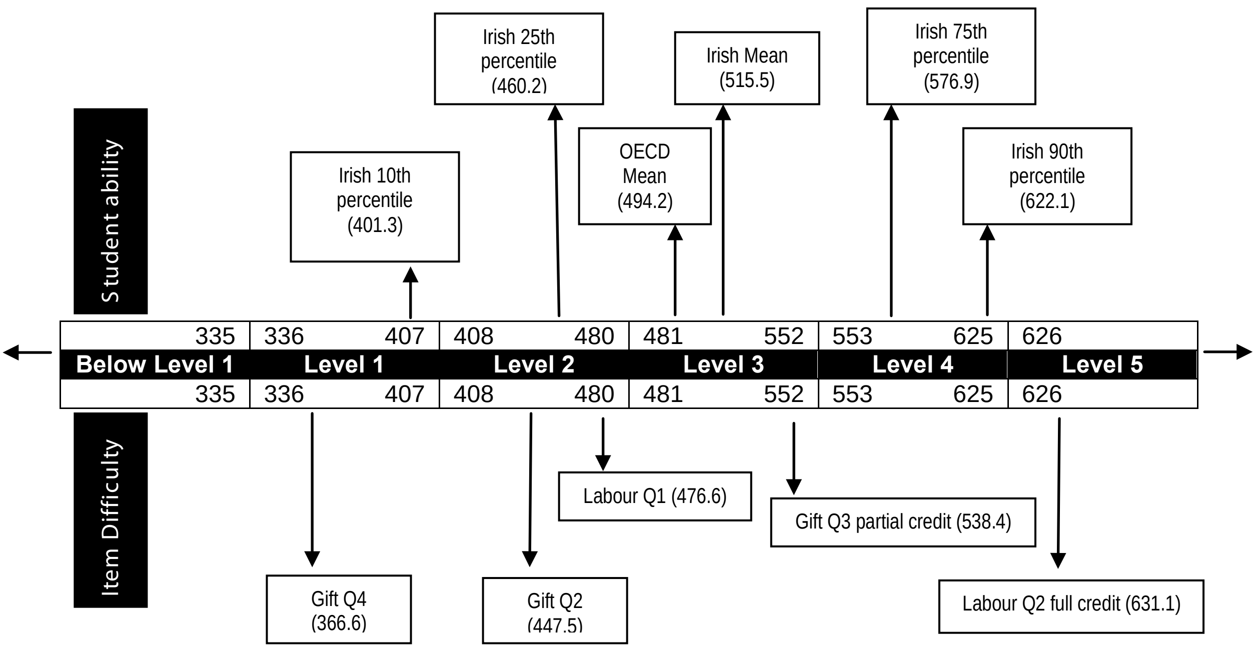
There is wide variation between countries in performance. For example, almost half (48.1%) of students attain Level 4 or 5 in Finland, and over 40% attain these levels in Korea, Canada, Australia, New Zealand, and Liechtenstein. In contrast, less than 10% of students attain Level 4 or 5 in Thailand, Serbia and Montenegro, Brazil, Mexico, Indonesia, and Tunisia. Less than 10% of students in Finland, Korea, and Canada are at or below Level 1; this again stands in sharp contrast to Mexico, Indonesia, and Tunisia, where over half of students in these countries score at or below the lowest level.

Table 3.21. *Descriptions of Proficiency Levels on the Reading Literacy Scale, and Percentages of Students Achieving at Each Level – Ireland and OECD*

		*Ireland**		*OECD***	
Level	*Brief Description – Students at this level are likely to be able to:*	*%*	*(SE)*	*%*	*(SE)*
Level 5 (greater than 625.6)	Complete the most complex PISA reading tasks, including managing information that is difficult to locate in complex texts; evaluate texts critically; and draw on specialised information.	9.3	(0.71)	8.3	(0.12)
Level 4 (552.9 to 625.6)	Complete difficult reading tasks, such as locating embedded information, constructing meaning from nuances of language, and critically evaluating a text.	26.2	(1.19)	21.3	(0.18)
Level 3 (480.2 to 552.9)	Complete reading tasks of moderate complexity, including locating multiple pieces of information, drawing links between different parts of a text, and relating text information to familiar everyday knowledge.	32.4	(1.26)	28.7	(0.19)
Level 2 (407.5 to 480.2)	Complete basic reading tasks, including locating one or more pieces of information which may require meeting multiple criteria, making low-level inferences of various types, and using some outside knowledge to understand text.	21.2	(1.20)	22.8	(0.18)
Level 1 (334.8 to 407.5)	Complete the most basic PISA reading tasks, such as locating a single piece of information, identifying the main theme of a text, and making a simple connection with everyday knowledge.	8.3	(0.66)	12.4	(0.16)
Below Level 1 (below 334.8)	Has a less than .50 chance of responding correctly to Level 1 tasks. Reading abilities not assessed by PISA.	2.7	(0.48)	6.7	(0.14)
	Total	100.0		100.0	

*N (Ireland) = 3880. **Denotes OECD average percent.
Note. Students at a level have at least a 50% chance of correctly answering all items at that level.

Figure 3.5. *The PISA 2003 Reading Literacy Scale: Cut-points for Proficiency Levels, Scores of Students in Ireland at Key Markers, and Difficulties of Selected Items*



Although New Zealand has a mean score does not differ significantly from that of Ireland, it has proportionately more students at or below Level 1 (14.5% compared to 11.0%), and more students at Levels 4 and 5 (40.6% compared to 35.5%), indicating a wider spread in achievement in New Zealand than in Ireland. The performance of Macao-China is also noteworthy: although its mean performance is significantly lower than that of Ireland, the proportion at or below Level 1 (9.7%) is slightly lower.

Table 3.22. Percentage of Students at Each Proficiency Level on the Reading Literacy Scale – OECD and Partner Countries

	< Level 1 (below 334.8)		*Level 1 (334.8 to 407.5)*		*Level 2 (407.5 to 480.2)*		*Level 3 (480.2 to 552.9)*		*Level 4 (552.9 to 625.6)*		*Level 5 (above 625.6)*	
	%	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*
Finland	1.1	(0.18)	4.6	(0.43)	14.6	(0.60)	31.7	(0.79)	33.4	(0.72)	14.7	(0.71)
Korea	1.4	(0.27)	5.4	(0.57)	16.8	(0.96)	33.5	(1.21)	30.8	(1.06)	12.2	(1.07)
Canada	2.3	(0.21)	7.3	(0.52)	18.3	(0.57)	31.0	(0.69)	28.6	(0.64)	12.6	(0.53)
Australia	3.6	(0.35)	8.2	(0.42)	18.3	(0.59)	28.4	(0.79)	26.9	(0.78)	14.6	(0.72)
Liechtenstein	2.5	(1.00)	7.9	(1.68)	18.7	(3.16)	30.3	(2.86)	27.6	(2.75)	13.0	(2.51)
New Zealand	4.8	(0.47)	9.7	(0.64)	18.5	(0.94)	26.3	(0.91)	24.3	(0.90)	16.3	(0.75)
Ireland	**2.7**	**(0.48)**	**8.3**	**(0.66)**	**21.2**	**(1.20)**	**32.4**	**(1.26)**	**26.2**	**(1.19)**	**9.3**	**(0.71)**
Sweden	3.9	(0.50)	9.4	(0.67)	20.7	(0.97)	29.9	(1.50)	24.8	(1.20)	11.4	(0.66)
Netherlands	2.1	(0.54)	9.4	(0.94)	23.4	(1.14)	30.7	(1.35)	25.6	(1.14)	8.8	(0.69)
Hong Kong-China	3.4	(0.67)	8.6	(0.77)	20.0	(1.04)	35.1	(1.24)	27.1	(1.19)	5.7	(0.51)
Belgium	7.8	(0.65)	10.0	(0.58)	18.2	(0.61)	26.0	(0.79)	25.4	(0.77)	12.5	(0.53)
Norway	6.4	(0.57)	11.8	(0.75)	21.4	(1.21)	29.0	(1.01)	21.5	(0.84)	10.0	(0.71)
Switzerland	5.4	(0.54)	11.3	(0.72)	22.7	(1.06)	30.9	(1.45)	21.9	(0.94)	7.9	(0.77)
Japan	7.4	(0.78)	11.6	(0.78)	20.9	(1.04)	27.2	(1.08)	23.2	(1.06)	9.7	(0.95)
Macao-China	1.0	(0.30)	8.7	(1.27)	27.8	(1.88)	41.4	(1.71)	19.4	(1.62)	1.7	(0.53)
Poland	5.3	(0.55)	11.5	(0.66)	24.4	(0.85)	30.0	(0.87)	20.7	(0.94)	8.0	(0.59)
France	6.3	(0.74)	11.2	(0.68)	22.8	(0.80)	29.7	(1.10)	22.5	(0.87)	7.4	(0.58)
United States	6.5	(0.67)	12.9	(0.92)	22.7	(1.06)	27.8	(1.05)	20.8	(0.90)	9.3	(0.66)
Denmark	4.6	(0.62)	11.9	(0.70)	24.9	(1.08)	33.4	(1.15)	20.0	(0.96)	5.2	(0.49)
Iceland	6.7	(0.61)	11.8	(0.69)	23.9	(0.85)	29.7	(1.00)	20.9	(0.83)	7.1	(0.59)
Germany	9.3	(0.78)	13.0	(0.87)	19.8	(0.84)	26.3	(0.80)	21.9	(0.96)	9.6	(0.64)
Austria	7.3	(0.79)	13.4	(0.97)	22.6	(1.02)	27.4	(1.03)	21.0	(1.00)	8.3	(0.79)
Latvia	5.0	(0.65)	13.0	(1.01)	25.6	(1.23)	30.8	(1.29)	19.5	(1.26)	6.0	(0.74)
Czech Republic	6.5	(0.93)	12.9	(0.85)	24.7	(0.98)	30.3	(1.28)	19.3	(1.08)	6.4	(0.60)
Hungary	6.1	(0.65)	14.4	(0.91)	26.7	(0.94)	30.2	(1.07)	17.6	(1.12)	4.9	(0.63)
Spain	7.4	(0.65)	13.7	(0.65)	26.1	(0.71)	29.6	(0.82)	18.2	(0.86)	5.0	(0.47)
Luxembourg	8.7	(0.39)	14.0	(0.66)	24.2	(0.70)	28.7	(0.99)	19.1	(0.85)	5.2	(0.37)
Portugal	7.6	(0.89)	14.4	(0.89)	25.9	(1.01)	30.5	(1.09)	17.9	(1.01)	3.8	(0.54)
Italy	9.1	(0.94)	14.8	(0.77)	24.9	(0.76)	28.3	(0.98)	17.8	(0.75)	5.2	(0.31)
Greece	10.2	(0.85)	15.0	(0.78)	25.0	(1.20)	27.3	(1.09)	16.8	(1.19)	5.7	(0.69)
Slovak Republic	8.0	(0.80)	16.9	(0.97)	28.4	(1.05)	27.7	(1.07)	15.4	(0.73)	3.5	(0.35)
Russian Federation	12.8	(1.12)	21.3	(0.97)	30.4	(1.02)	24.5	(1.11)	9.3	(0.76)	1.7	(0.29)
Turkey	12.5	(1.24)	24.3	(1.51)	30.9	(1.44)	20.8	(1.44)	7.7	(1.14)	3.8	(1.19)
Uruguay	20.2	(1.04)	19.6	(0.83)	23.9	(0.84)	19.8	(0.89)	11.2	(0.78)	5.3	(0.66)
Thailand	13.5	(0.99)	30.5	(1.16)	34.3	(1.01)	17.0	(0.90)	4.1	(0.55)	0.5	(0.15)
Serbia and Montenegro	17.1	(1.14)	29.6	(1.30)	33.3	(1.09)	16.4	(1.14)	3.5	(0.59)	0.2	(0.11)
Brazil	26.9	(1.58)	23.1	(1.17)	25.2	(0.96)	16.5	(0.99)	6.3	(0.66)	1.9	(0.49)
Mexico	24.9	(1.52)	27.1	(1.18)	27.5	(1.04)	15.6	(0.97)	4.3	(0.59)	0.5	(0.14)
Indonesia	26.0	(1.54)	37.2	(1.15)	27.3	(1.13)	8.2	(0.87)	1.2	(0.28)	0.1	(0.05)
Tunisia	33.7	(1.30)	29.0	(0.90)	23.6	(0.90)	10.9	(0.76)	2.5	(0.43)	0.3	(0.11)
OECD Total	8.1	(0.27)	13.6	(0.33)	22.9	(0.40)	27.2	(0.35)	20.1	(0.31)	8.1	(0.22)
OECD Average	6.7	(0.14)	12.4	(0.16)	22.8	(0.18)	28.7	(0.19)	21.3	(0.18)	8.3	(0.12)

Note. OECD countries are in regular font, partner countries are in *italics* . Countries are ordered in descending order of mean score on the combined reading literacy scale.

Variation in Performance on the Reading Literacy Scale

In the previous section, the performance of students was described in terms of specified levels of knowledge and skills relative to absolute benchmarks (the proficiency levels). This section considers the relative dispersion of scores by examining the gap between the highest and lowest performing students. Small gaps are indicative of higher levels of similarity in outcomes. The focus of this section is on the 10th and 90th percentiles. Table 3.23 also shows the scores at the 5th, 25th, 75th, and 95th percentiles.

Table 3.23. Scores at the 5th, 10th, 25th, 75th, 90th, and 95th Percentiles on the Reading Literacy Scale – OECD and Partner Countries

	5th		*10th*		*25th*		*75th*		*90th*		*95th*	
	Score	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*
Finland	399.7	(4.82)	437.0	(3.14)	493.6	(2.42)	598.8	(1.67)	641.1	(2.16)	665.8	(2.48)
Korea	393.3	(5.99)	427.9	(5.25)	484.2	(4.07)	590.0	(2.82)	633.9	(4.13)	659.6	(5.02)
Canada	372.5	(3.12)	410.0	(3.14)	472.2	(2.30)	589.8	(2.05)	635.6	(2.13)	663.0	(2.53)
Australia	352.1	(4.84)	394.9	(3.64)	464.3	(3.01)	593.5	(2.51)	644.0	(2.66)	672.8	(3.10)
Liechtenstein	365.4	(15.05)	404.8	(11.66)	467.1	(9.12)	588.4	(5.73)	636.2	(11.80)	661.0	(14.25)
New Zealand	337.6	(6.20)	380.9	(4.42)	452.8	(3.53)	595.9	(2.83)	652.0	(2.94)	681.7	(3.43)
Ireland	**363.5**	**(7.34)**	**401.3**	**(4.64)**	**460.2**	**(3.80)**	**576.9**	**(2.77)**	**622.1**	**(3.00)**	**647.4**	**(3.27)**
Sweden	348.6	(5.99)	389.7	(4.33)	453.1	(3.39)	581.6	(2.90)	631.3	(2.86)	660.2	(3.63)
Netherlands	368.7	(6.42)	400.4	(5.16)	454.1	(4.52)	576.1	(3.20)	620.5	(2.89)	644.9	(4.17)
Hong Kong-China	354.9	(9.87)	396.6	(6.65)	460.8	(5.10)	568.9	(2.77)	608.0	(2.94)	629.7	(2.99)
Belgium	299.9	(8.39)	354.9	(6.58)	439.9	(4.23)	587.0	(2.11)	635.1	(2.15)	661.9	(2.61)
Norway	321.1	(6.08)	363.9	(4.71)	434.3	(3.83)	570.6	(3.59)	625.1	(3.89)	656.1	(3.93)
Switzerland	330.3	(5.84)	373.4	(5.64)	438.5	(4.51)	565.3	(3.67)	615.0	(3.93)	643.4	(5.04)
Japan	310.1	(7.32)	354.7	(6.51)	431.2	(5.35)	574.1	(3.66)	623.6	(4.76)	652.4	(4.74)
Macao-China	381.4	(6.23)	408.7	(5.09)	454.6	(3.50)	544.2	(4.44)	582.7	(3.71)	601.4	(4.31)
Poland	330.5	(6.29)	373.6	(4.96)	436.4	(3.56)	562.6	(3.08)	615.8	(3.45)	645.2	(4.42)
France	320.4	(7.74)	367.0	(7.00)	436.0	(3.98)	565.4	(2.82)	613.5	(2.71)	640.6	(3.28)
United States	319.0	(6.56)	360.7	(5.25)	429.0	(4.15)	567.9	(3.59)	622.1	(3.55)	651.4	(4.49)
Denmark	337.9	(6.57)	376.2	(4.56)	437.8	(3.98)	553.4	(2.97)	600.5	(2.68)	626.7	(3.92)
Iceland	316.4	(6.38)	362.3	(4.78)	430.7	(2.31)	560.3	(2.24)	611.6	(2.84)	639.5	(3.64)
Germany	294.7	(5.95)	340.7	(6.78)	418.5	(5.61)	571.7	(3.39)	623.5	(3.19)	652.0	(3.86)
Austria	313.1	(7.49)	353.6	(6.34)	423.5	(4.92)	565.1	(4.19)	617.4	(3.66)	645.9	(4.69)
Latvia	334.5	(6.37)	371.9	(5.30)	431.0	(4.92)	554.3	(3.50)	602.7	(4.62)	632.2	(4.62)
Czech Republic	320.2	(9.47)	361.9	(6.87)	427.8	(4.69)	555.3	(4.00)	606.9	(3.84)	636.2	(4.00)
Hungary	323.9	(5.95)	361.2	(4.24)	422.0	(3.30)	546.5	(3.27)	597.1	(3.37)	624.9	(4.97)
Spain	312.8	(5.77)	353.6	(4.87)	420.5	(3.37)	548.2	(2.79)	597.0	(2.76)	625.1	(3.09)
Luxembourg	301.7	(3.84)	343.7	(2.87)	415.7	(2.78)	551.4	(1.86)	600.6	(2.06)	627.1	(2.72)
Portugal	311.0	(6.61)	351.4	(7.07)	417.6	(5.19)	543.8	(3.48)	591.6	(3.54)	616.6	(3.94)
Italy	294.9	(8.59)	341.3	(6.76)	411.4	(4.39)	547.0	(2.53)	598.4	(2.07)	626.9	(2.56)
Greece	287.8	(6.16)	333.2	(6.22)	406.3	(5.18)	545.9	(4.39)	599.4	(4.37)	630.9	(5.39)
Slovak Republic	310.0	(5.66)	347.8	(5.85)	407.9	(4.58)	535.3	(3.16)	586.6	(2.97)	613.0	(3.45)
Russian Federation	281.2	(6.87)	318.6	(6.11)	381.0	(5.43)	506.5	(3.86)	558.3	(4.43)	588.1	(4.66)
Turkey	290.8	(6.14)	323.7	(5.33)	377.0	(5.69)	499.9	(6.61)	562.2	(11.42)	607.9	(19.42)
Uruguay	223.9	(5.80)	272.2	(5.99)	355.3	(4.41)	518.3	(4.36)	587.4	(4.52)	628.2	(6.08)
Thailand	293.2	(4.88)	321.7	(3.44)	366.0	(3.10)	471.6	(3.56)	520.3	(4.54)	549.7	(5.26)
Serbia and Montenegro	274.2	(4.99)	305.8	(4.65)	357.9	(4.01)	466.9	(4.00)	515.7	(4.81)	541.7	(5.86)
Brazil	213.7	(7.32)	255.8	(7.48)	328.0	(5.49)	479.3	(5.09)	542.3	(5.22)	580.8	(6.94)
Mexico	238.1	(6.10)	273.9	(5.51)	335.1	(4.95)	466.6	(4.35)	520.6	(6.07)	551.7	(5.52)
Indonesia	254.2	(5.35)	281.8	(4.94)	332.0	(3.74)	432.9	(4.01)	477.6	(4.59)	505.6	(6.08)
Tunisia	216.0	(4.73)	250.6	(3.83)	309.7	(3.18)	441.1	(3.48)	496.5	(4.26)	529.6	(5.53)
OECD Total	**304.8**	**(2.25)**	**348.6**	**(2.17)**	**420.1**	**(1.82)**	**562.1**	**(1.25)**	**615.8**	**(1.24)**	**645.7**	**(1.30)**
OECD Average	**317.6**	**(1.43)**	**360.8**	**(1.29)**	**430.1**	**(0.99)**	**565.3**	**(0.63)**	**616.9**	**(0.64)**	**646.2**	**(0.70)**

Note. OECD countries are in regular font, partner countries are in *italics*. Countries are ordered in descending order of mean score on the combined reading literacy scale.

The OECD average difference between the 10th and 90th percentiles is 256.1 points which illustrates the wide variation in performance amongst students. This is larger than the difference between the mean scores of the highest and lowest scoring countries (168.9 points) and demonstrates that differences within countries are larger than differences among countries. The difference between the 10th and 90th percentiles in Ireland is 220.8, indicating a comparatively narrow spread in achievement. This stands in contrast to New Zealand (with a difference of 271.1 points) and Sweden (241.6 points), countries with similar mean scores to that of Ireland. In Macao-China, the difference between the top and bottom 10% of students is just 174.0 score points. Finland and Korea are also noteworthy, achieving both high mean achievement scores (543.5 and 534.1, respectively) and comparatively small differences between high and low performers (204.1 and 206.0 points, respectively).

Irish students at the 90th percentile achieved a score of 622.1, which is just over 5 points above the OECD average of 616.9. In contrast, the score for Irish students at the 10th percentile (401.3), is 40.5 score points higher than the OECD average of 360.5. This demonstrates that, like mathematics, Ireland's high average performance is due to comparatively high achievements at the lower end of the scale. Comparing these scores with those of Australia and Sweden (countries with mean scores similar to Ireland), it can be seen that Irish students are doing comparatively better at the 10th percentile and comparatively less well at the 90th.

ACHIEVEMENT IN SCIENCE

The science assessment measured students' ability to describe, explain and predict scientific phenomena, to understand and interpret scientific evidence, and to draw conclusions. Three major areas were assessed: life and health, Earth and environment, and technology. This section addresses the interpretation of scores, the overall performance (mean scores) of students in Ireland and in other countries, and the performance of Irish students at the 10th and 90th percentiles. Examples of contexts and items used to assess students' knowledge of science may be viewed at http://www.erc.ie/pisa/, which also includes scale score values for selected items. As science was a minor assessment domain in PISA 2000 and PISA 2003, there was insufficient information available to develop proficiency levels similar to those developed for mathematics and reading literacy.[21]

Interpreting Scores on the Science Scale

Although it was not possible to develop proficiency levels, it was possible to generate a description of the skills associated with different points along the science scale using procedures similar to those used to describe the skills associated with proficiency levels for mathematics and reading literacy (see Inset 3.3). Towards the top of the science scale (around 690 points), students were likely to complete the following tasks:

- create or use conceptual models to make predictions or give explanations;
- analyse scientific investigations in relation to experimental design;
- use data as evidence to evaluate alternative viewpoints; and
- communicate scientific arguments and descriptions in detail.

[21] This will be possible in 2006, however, when science becomes the major assessment domain of PISA.

At an intermediate point on the scale (around 550 points), students were likely to:

- use scientific concepts to make predictions or give explanations;
- recognise questions that can be answered by scientific investigation; and
- select relevant information from competing data or chains of reasoning in drawing or evaluating conclusions.

Towards the lower end of the scale (around 400 score points), students were likely to:

- recall simple scientific factual knowledge; and
- use common science knowledge in drawing or evaluating conclusions.

Mean Scores on the Science Scale

Ireland achieved a mean score of 505.4 on the science scale (Table 3.24). Although just under 6 points higher than the OECD average of 499.6, the difference is nonetheless statistically significant. Ireland's ranking in science is 16th of 40 countries (95% confidence interval of Ireland's ranking = 12th to 20th), and 13th out of 29 OECD countries (95% confidence interval of Ireland's ranking = 9th to 16th). Countries with mean scores that are significantly higher than Ireland's include Finland, Japan, Hong Kong-China, Korea, Australia and New Zealand. The mean scores of 11 countries (eight OECD countries) are significantly higher than that of Ireland. Ireland's mean score does not differ significantly from the mean scores of eight countries (all OECD countries) and is significantly higher than those of 20 countries (11 OECD countries). The highest scoring countries, Finland and Japan, have mean scores that are approximately 40.0 points higher than that of Ireland. The difference between the highest and lowest scoring country is 163.5 points. As in the case of mathematics and reading literacy, performance is more clustered at the upper end of the distribution of countries. For example, just 23.5 points separate the top seven performing countries, while 51.7 points separate the seven lowest performing countries.

The standard deviation for Ireland (93.0) is lower than the OECD average (105.5), indicating a comparatively narrow range of achievement, which was also observed for mathematics and reading. In contrast, the standard deviations in science for countries such as Belgium (107.4) and Germany (111.4) are high.

It is interesting to observe that all 11 countries that have significantly higher mean scores in science than Ireland also have significantly higher mean scores in combined mathematics. However, just three of these, Finland, Korea, and Canada, have a significantly higher score on reading literacy. Four additional countries, with significantly higher mean scores on combined mathematics than Ireland, have mean scores in science that are not significantly different from that of Ireland, or are significantly lower. Thus, in relative terms, Ireland's performance in science is somewhat better than in combined mathematics, but poorer than in reading literacy.

Table 3.24. Mean Achievement Scores and Standard Deviations on the Science Scale – OECD and Partner Countries

	Pop	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*		*Pop*	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*
Finland	⊠	548.2	(1.92)	90.8	(1.05)	▲	United States	⊠	491.3	(3.08)	101.6	(1.34)	▼
Japan	⊠	547.6	(4.14)	109.4	(2.71)	▲	Austria	⊠	491.0	(3.44)	97.0	(1.51)	▼
Hong Kong - Ch	⊠	539.5	(4.26)	94.1	(2.78)	▲	*Russian Fed*	⊠	489.3	(4.14)	99.8	(1.49)	▼
Korea	⊠	538.4	(3.54)	100.5	(2.16)	▲	*Latvia*	⊠	489.1	(3.89)	92.7	(1.48)	▼
Liechtenstein	□	525.2	(4.33)	103.5	(4.35)	▲	Spain	⊠	487.1	(2.61)	100.2	(1.51)	▼
Australia	⊠	525.1	(2.10)	101.8	(1.53)	▲	Italy	⊠	486.5	(3.13)	107.8	(2.02)	▼
Macao - Ch	□	524.7	(3.03)	87.9	(3.04)	▲	Norway	⊠	484.2	(2.87)	103.8	(1.82)	▼
Netherlands	⊠	524.4	(3.15)	98.5	(2.17)	▲	Luxembourg	⊠	482.8	(1.50)	102.8	(1.13)	▼
Czech Rep	⊠	523.3	(3.38)	100.6	(1.69)	▲	Greece	⊠	481.0	(3.82)	100.6	(1.65)	▼
New Zealand	⊠	520.9	(2.35)	104.0	(1.39)	▲	Denmark	⊠	475.2	(2.97)	101.8	(1.66)	▼
Canada	⊠	518.7	(2.02)	99.1	(1.05)	▲	Portugal	⊠	467.7	(3.46)	93.4	(1.74)	▼
Switzerland	⊠	513.0	(3.69)	107.5	(1.85)	▲	*Uruguay*	×	438.4	(2.90)	109.1	(1.83)	▼
France	⊠	511.2	(2.99)	110.8	(2.17)	▲	*Serbia and Monte.*	⊠	436.4	(3.50)	82.7	(1.61)	▼
Belgium	⊠	508.8	(2.48)	107.4	(1.82)	▲	*Turkey*	×	434.2	(5.89)	95.9	(4.71)	▼
Sweden	⊠	506.1	(2.72)	106.8	(1.81)	▲	*Thailand*	□	429.1	(2.70)	81.3	(1.60)	▼
Ireland	⊠	**505.4**	**(2.69)**	**93.0**	**(1.33)**	▲	Mexico	×	404.9	(3.49)	86.7	(2.20)	▼
Hungary	⊠	503.3	(2.77)	97.3	(1.99)	o	*Indonesia*	×	395.0	(3.21)	68.0	(1.87)	▼
Germany	⊠	502.3	(3.64)	111.4	(2.15)	o	*Brazil*	×	389.6	(4.35)	98.3	(2.63)	▼
Poland	⊠	497.8	(2.86)	102.4	(1.38)	o	*Tunisia*	⊠	384.7	(2.56)	87.3	(1.84)	▼
Slovak Rep	⊠	494.9	(3.71)	102.2	(3.11)	o	**OECD Total**		**495.7**	**(1.07)**	**109.0**	**(0.69)**	
Iceland	⊠	494.7	(1.47)	95.6	(1.43)	▼	**OECD Average**		**499.6**	**(0.60)**	**105.5**	**(0.37)**	

⊠ >90% of 15-year olds enrolled
□ 75-90% of 15-year olds enrolled
× 50-75% of 15-year olds enrolled

■ Mean significantly higher than Ireland
□ Mean not significantly different from Ireland
▒ Mean significantly lower than Ireland

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

Note. OECD countries are in regular font; partner countries are in italics. SD = Standard deviation; SE = Standard error.
The column "Pop" is an indicator of the percent of the 15-year-old population enrolled in schools in each country and is based on Column 15 of Table A3.1 in OECD (2004b).
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.

When country mean scores are compared using the nonparametric maximum likelihood estimation method (see Table 3.25 and Figure 3.6, as well as Inset 3.2 for information on interpretation), countries are categorised into six distinct performance groupings. Ireland's performance is indistinguishable from that of Belgium, Sweden, Hungary, and Germany in the third grouping, where the mean is 505.0.

Table 3.25. Seven-point NPML Probability Distribution and Posterior Probabilities for the Science Scale – OECD and Partner Countries

	Rank	*Masspoints*						
PISA score		*390.6*	*403.3*	*434.6*	*485.6*	*505.0*	*522.8*	*542.5*
Proportion		*0.082*	*0.030*	*0.146*	*0.300*	*0.160*	*0.164*	*0.119*
Finland	1	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Japan	2	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Hong Kong-China	3	0.00	0.00	0.00	0.00	0.00	0.02	0.98
Korea	4	0.00	0.00	0.00	0.00	0.00	0.04	0.96
Liechtenstein	5	0.00	0.00	0.00	0.00	0.00	0.99	0.01
Australia	6	0.00	0.00	0.00	0.00	0.00	0.99	0.01
Macao-China	7	0.00	0.00	0.00	0.00	0.00	0.99	0.00
Netherlands	8	0.00	0.00	0.00	0.00	0.00	0.99	0.00
Czech Republic	9	0.00	0.00	0.00	0.00	0.00	0.99	0.00
New Zealand	10	0.00	0.00	0.00	0.00	0.02	0.98	0.00
Canada	11	0.00	0.00	0.00	0.00	0.06	0.94	0.00
Switzerland	12	0.00	0.00	0.00	0.00	0.62	0.38	0.00
France	13	0.00	0.00	0.00	0.00	0.82	0.18	0.00
Belgium	14	0.00	0.00	0.00	0.00	0.95	0.05	0.00
Sweden	15	0.00	0.00	0.00	0.00	0.99	0.01	0.00
Ireland	**16**	**0.00**	**0.00**	**0.00**	**0.00**	**0.99**	**0.01**	**0.00**
Hungary	17	0.00	0.00	0.00	0.01	0.99	0.00	0.00
Germany	18	0.00	0.00	0.00	0.02	0.98	0.00	0.00
Poland	19	0.00	0.00	0.00	0.28	0.72	0.00	0.00
Slovak Republic	20	0.00	0.00	0.00	0.71	0.29	0.00	0.00
Iceland	21	0.00	0.00	0.00	0.72	0.28	0.00	0.00
United States	22	0.00	0.00	0.00	0.96	0.04	0.00	0.00
Austria	23	0.00	0.00	0.00	0.96	0.04	0.00	0.00
Russian Federation	24	0.00	0.00	0.00	0.99	0.01	0.00	0.00
Latvia	25	0.00	0.00	0.00	0.99	0.01	0.00	0.00
Spain	26	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Italy	27	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Norway	28	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Luxembourg	29	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Greece	30	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Denmark	31	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Portugal	32	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Uruguay	33	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Serbia and Montenegro	34	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Turkey	35	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Thailand	36	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Mexico	37	0.09	0.91	0.00	0.00	0.00	0.00	0.00
Indonesia	38	0.86	0.14	0.00	0.00	0.00	0.00	0.00
Brazil	39	0.98	0.02	0.00	0.00	0.00	0.00	0.00
Tunisia	40	1.00	0.00	0.00	0.00	0.00	0.00	0.00

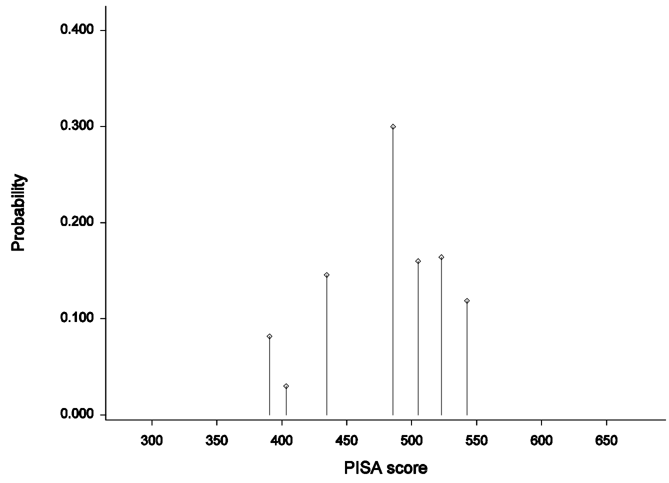
■ Probability of masspoint grouping is .95 or more; high degree of confidence

▒ Probability of masspoint grouping is .90 to .95; borderline degree of confidence

□ Probability of masspoint grouping is less than .90; low degree of confidence

OECD countries are in regular font; partner countries are in *italics* .

Figure 3.6. Plot of the NPML Probability Distribution for Science Country Scores



Variation in Performance on the Science Scale

In this section, the performance of students at the 10th and 90th percentiles on the science scale is considered. The smaller the difference, the less dispersed student achievement is within a country. Table 3.26 shows the scores at these percentile points, as well as the 5th, 25th, 75th and 95th percentiles for the 40 participating countries that satisfied PISA sampling requirements. Figure 3.7 shows the relationship between Irish student proficiency and item difficulty for science using a selection of the sample items available at http://www.erc.ie/pisa/ and a selection of key performance benchmarks. For example, the second item (Q2) from the unit 'Semmelweis' has an item difficulty (494.9) that is a little below the Irish student mean (505.4). The partial credit version of the first item from the unit 'Semmelweis' has an item difficulty (625.7) which is very close to the Irish 90th percentile (624.7).

Figure 3.7. The PISA 2003 Science Scale: Cut-points for Key OECD Percentile Intervals, Scores of Students in Ireland at Key Markers, and Difficulties of Selected Items

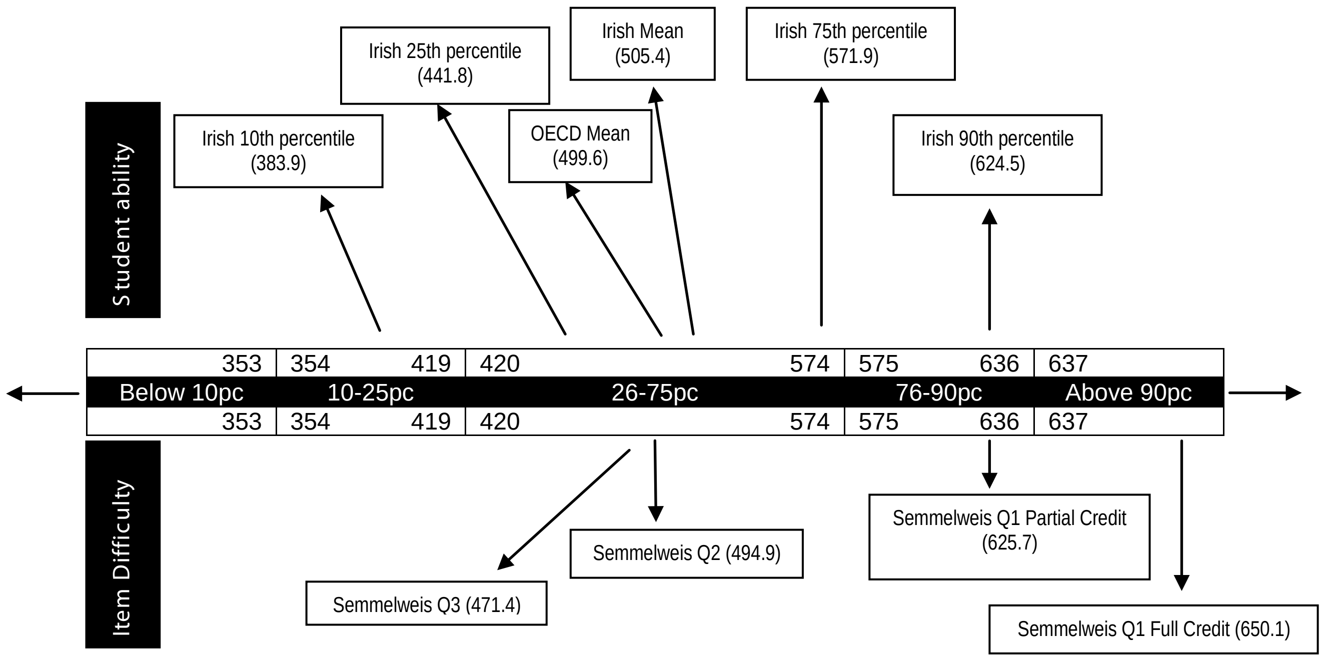


Table 3.26. *Scores at the 5th, 10th, 25th, 75th, 90th, and 95th Percentiles on the Science Scale – OECD and Partner Countries*

	5th		*10th*		*25th*		*75th*		*90th*		*95th*	
	Score	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*	*Score*	*(SE)*
Finland	392.7	(3.46)	429.3	(2.56)	488.4	(2.80)	611.1	(2.22)	661.7	(2.90)	690.6	(3.54)
Japan	357.0	(6.96)	402.3	(6.04)	475.1	(6.10)	624.3	(4.20)	681.5	(6.00)	715.3	(7.94)
Hong Kong - China	372.9	(9.84)	412.4	(8.61)	478.3	(6.92)	607.6	(3.47)	653.4	(3.88)	679.8	(4.28)
Korea	364.8	(6.27)	404.9	(4.97)	472.5	(4.76)	608.7	(4.26)	662.8	(4.66)	695.0	(5.83)
Liechtenstein	350.9	(17.29)	388.7	(8.65)	450.4	(5.74)	598.0	(9.11)	659.1	(10.38)	690.2	(13.46)
Australia	351.3	(4.16)	390.7	(3.40)	457.3	(3.13)	596.4	(2.66)	652.4	(2.92)	685.9	(3.74)
Macao - China	375.3	(7.90)	409.7	(7.72)	464.9	(5.29)	587.2	(3.99)	635.1	(6.25)	663.0	(9.50)
Netherlands	362.8	(6.63)	393.9	(5.56)	451.4	(5.35)	598.5	(4.02)	653.3	(4.07)	682.0	(4.27)
Czech Republic	356.5	(5.79)	391.4	(4.26)	452.9	(4.23)	594.4	(3.95)	652.0	(4.66)	686.4	(4.53)
New Zealand	347.5	(3.94)	381.6	(4.10)	447.9	(3.92)	596.0	(3.34)	653.2	(3.91)	687.0	(3.20)
Canada	351.6	(3.92)	388.8	(3.35)	452.4	(2.65)	588.3	(2.43)	643.9	(3.03)	675.9	(2.90)
Switzerland	327.9	(5.83)	369.4	(4.63)	440.4	(4.47)	588.3	(4.65)	647.9	(5.90)	683.3	(6.81)
France	320.8	(6.74)	363.4	(5.51)	435.0	(4.40)	591.4	(3.37)	650.9	(3.21)	681.9	(4.54)
Belgium	319.5	(6.15)	363.8	(5.02)	436.2	(3.77)	588.4	(2.40)	640.2	(2.51)	668.5	(2.56)
Sweden	327.4	(6.53)	368.4	(4.03)	434.6	(3.47)	581.1	(3.96)	642.2	(3.99)	673.2	(4.76)
Ireland	**347.9**	**(6.14)**	**383.9**	**(4.78)**	**441.8**	**(3.68)**	**571.9**	**(3.00)**	**624.5**	**(3.32)**	**652.3**	**(3.44)**
Hungary	339.6	(5.92)	375.4	(4.06)	436.8	(3.08)	572.1	(3.85)	627.6	(5.53)	657.7	(4.61)
Germany	306.7	(7.15)	350.5	(5.55)	426.7	(5.83)	583.7	(3.95)	640.3	(3.65)	671.7	(3.53)
Poland	333.2	(5.30)	366.6	(3.46)	426.2	(4.27)	569.5	(3.50)	630.1	(4.11)	666.1	(6.31)
Slovak Republic	331.5	(7.00)	367.1	(5.96)	427.5	(4.58)	565.7	(3.62)	624.9	(3.82)	656.8	(3.86)
Iceland	331.3	(5.87)	369.1	(3.97)	432.0	(2.80)	561.6	(2.69)	616.0	(3.61)	646.5	(3.61)
United States	321.7	(5.43)	358.7	(4.41)	420.2	(3.80)	564.1	(3.32)	622.3	(4.32)	654.3	(3.53)
Austria	327.1	(6.60)	362.6	(4.12)	423.1	(4.14)	561.2	(3.98)	614.7	(4.05)	644.4	(4.41)
Russian Federation	323.8	(5.55)	359.2	(5.42)	422.4	(4.79)	558.0	(4.49)	616.8	(4.04)	652.1	(4.95)
Latvia	336.1	(5.62)	369.9	(5.04)	425.3	(4.62)	553.1	(5.09)	608.6	(4.90)	641.8	(5.71)
Spain	317.7	(5.80)	355.4	(4.04)	420.7	(3.42)	557.0	(3.09)	612.8	(3.05)	644.2	(3.79)
Italy	302.7	(7.26)	344.4	(6.30)	414.6	(4.92)	562.6	(2.75)	621.5	(2.65)	655.7	(3.94)
Norway	311.7	(5.29)	349.5	(4.59)	414.1	(4.02)	557.1	(3.83)	615.8	(4.57)	651.2	(6.05)
Luxembourg	309.1	(4.25)	347.3	(2.64)	412.8	(2.85)	556.3	(2.44)	614.3	(3.13)	645.4	(2.89)
Greece	314.5	(5.78)	349.4	(4.97)	411.8	(4.53)	552.1	(4.02)	610.3	(4.63)	643.4	(4.94)
Denmark	305.7	(6.37)	343.1	(4.71)	407.3	(3.87)	546.9	(3.60)	604.8	(3.39)	637.9	(4.38)
Portugal	309.7	(5.95)	346.3	(6.16)	404.7	(5.02)	533.1	(3.38)	586.8	(3.68)	618.4	(4.45)
Uruguay	257.0	(3.94)	295.8	(4.40)	362.9	(3.99)	515.6	(4.49)	578.8	(5.00)	612.8	(5.28)
Serbia and Montenegro	305.0	(4.46)	332.4	(3.92)	379.7	(3.87)	491.9	(4.42)	544.5	(5.23)	575.6	(6.41)
Turkey	294.5	(4.96)	320.6	(4.75)	367.3	(4.85)	491.6	(8.35)	559.7	(12.83)	608.6	(19.98)
Thailand	302.8	(3.57)	329.3	(3.35)	373.4	(2.88)	480.5	(3.51)	536.8	(4.44)	570.9	(5.56)
Mexico	263.8	(5.12)	295.0	(4.75)	347.0	(3.49)	462.3	(4.18)	517.3	(5.28)	550.6	(6.76)
Indonesia	285.5	(4.45)	310.1	(3.98)	349.9	(2.98)	438.5	(3.82)	483.4	(5.47)	511.9	(6.18)
Brazil	234.6	(7.58)	267.9	(5.17)	323.4	(4.77)	452.2	(5.43)	519.9	(7.56)	559.9	(7.88)
Tunisia	243.7	(4.60)	273.8	(3.75)	325.2	(2.72)	443.6	(3.27)	498.2	(4.97)	529.7	(6.24)
OECD total	**315.5**	**(1.91)**	**353.2**	**(1.62)**	**418.8**	**(1.66)**	**573.9**	**(1.40)**	**635.7**	**(1.47)**	**670.2**	**(1.65)**
OECD average	**323.9**	**(1.23)**	**361.6**	**(1.14)**	**426.9**	**(1.02)**	**574.6**	**(0.80)**	**634.2**	**(0.91)**	**667.5**	**(1.01)**

Note. OECD countries are in regular font, partner countries are in *italics* . Countries are ordered in descending order of mean score on the science scale.

The OECD average difference between the 10th and 90th percentiles is 272.6 points. This is much larger than the difference between the highest and lowest country mean (163.5 points) and shows that differences amongst students across countries far exceed mean differences between countries. The difference for Ireland between the 10th and 90th percentiles (240.6 points) is somewhat smaller than the corresponding average difference for OECD countries. Finland and Hong Kong-China not only have high average performance but also a comparatively small difference between these percentiles (232.4 and 241.0 respectively). Macao-China is again notable in its good overall performance and small dispersion of achievement (difference between 10th and 90th percentiles = 225.4 points).

The score for Ireland at the 10th percentile was 383.9, which is 22.3 points higher than the OECD average of 361.6. This score ranks Ireland 11th out of 40 countries at this marker. The score for Ireland at the 90th percentile on the science scale is 624.5, giving a rank of 20th. This suggests, as with both mathematics and reading, that low achievers in Ireland are performing comparatively well, while high achievers are performing comparatively poorly.

PERFORMANCE IN PISA 2003 COMPARED TO 2000

In this section, the performances of students in PISA in 2000 and 2003 are compared for the 32 countries for which comparisons across the two assessments are possible. Comparisons are not possible for the Netherlands and the UK; the former did not meet required response rate standards in 2000, and the latter failed to do so in 2003. Comparisons are not possible for Luxembourg either since the testing conditions are not comparable (in 2000, students were not given a choice of the language of the test; in 2003, they were). For the remaining countries, comparisons are not made because they did not participate in PISA 2000. Inset 3.4 outlines some important considerations in interpreting the information presented in this section.

Inset 3.4. Interpreting Changes in Achievement Between PISA 2000 and PISA 2003

It is not possible to compare achievement on the combined mathematics scale in 2000 and 2003 since the 2000 assessment included items on only two of the four subscales. Performance is compared for the two subscales common to both PISA 2000 and PISA 2003 (Space & Shape, and Change & Relationships).

To make comparisons possible, achievement data in reading, mathematics and science for PISA 2003 were scaled using the item parameter estimates from PISA 2000 for the set of items common to both assessments. This entailed the introduction of some additional measurement error due to the equating of the two assessments which is taken into account in the significance tests used to compare performance in 2000 and 2003. Specifically, the additional measurement or equating error is incorporated into the calculation of significance tests using the following formula:

Equating error = (Mean2000 - Mean2003) / SQRT ($SE2000^2 + SE2003^2 + SELinking^2$).

The linking error for the Space & Shape subscale is 6.0, for the subscale Change & Relationships it is 4.8, for reading literacy it is 3.7, and for science it is 3.0.

Caution should be exercised in interpreting comparisons between 2000 and 2003. First, the data represent only two points in time; further assessments would be needed to establish a reliable picture of trends. Second, methodologies for establishing links between assessments over time are still evolving. Third, the changes in performance should be interpreted with reference to any major policy or structural changes introduced to countries’ education systems.

Performance on the Mathematics Space and Shape Subscale in 2000 and 2003

Table 3.27 provides comparative data on the mean performance of countries on the Space & Shape subscale, as well as a comparison of performance at a number of percentile points (5th, 10th, 25th, 75th, 90th, 95th).

There are no significant differences between the OECD mean scores, or OECD average scores at the six percentile points, between 2000 and 2003. However, there has been a significant increase in the mean score in eight countries (four OECD countries), and a significant decrease in two OECD countries.

Changes in performance at the percentile points are also found in 18 countries. Significant increases and decreases are associated with both high- and low-performing countries. For some countries (e.g., the Czech Republic), there has been a significant increase in the achievements of average and lower performers; in others (e.g., Belgium), the increase is for average and higher performers. There is no change at any point in the performance scale for Irish students.

Performance on the Mathematics Change and Relationships Subscale in 2000 and 2003

Table 3.28 compares the mean performance of countries, and performance at a number of percentile points, for the Change & Relationships subscale in 2000 and 2003.

The OECD country average score is significantly higher in 2003 than in 2000, while significant increases are also observed at five of the six percentile points of interest. In 13 countries (10 OECD countries) there is a significant increase in mean performance, in one partner country (Thailand) a significant decrease, and no change in 18 countries (15 OECD countries), including Ireland.

Canada stands out as having a significant increase in mean score, and at each key percentile point. Significant increases again are associated with both high- and low-performing countries.

Performance on the Reading Literacy Scale in 2000 and 2003

Table 3.29 provides comparative data on countries' mean performance, and on country performance at a number of percentile points, for the reading literacy scale in 2000 and 2003. The OECD average scores at all seven percentile points are similar in 2000 and 2003. Three countries registered a significant increase in mean performance and nine countries (including Ireland) a significant decrease. There was no change in 20 countries.

There was a significant difference of 11.2 points between Ireland's mean score in 2000 (526.7) and 2003 (515.5). The performance of Irish students at the 75th, 90th, and 95th percentiles was also significantly lower in 2003 than in 2000. A broadly similar pattern of differences can be observed for Hong Kong-China.

Table 3.27. Comparison of Mean Performance and Performance at Key Percentiles on the Space and Shape Subscale, 2000 and 2003 – OECD and Partner Countries

	Point on the scale being compared						
	5th	*10th*	*25th*	*Mean*	*75th*	*90th*	*95th*
Hong Kong - China	o	o	o	o	▲	o	o
Japan	o	o	o	o	o	o	o
Korea	o	o	o	o	o	o	o
Switzerland	o	o	o	o	o	o	o
Finland	▲	o	o	o	o	o	o
Liechtenstein	o	o	o	o	o	o	o
Belgium	o	o	▲	▲	▲	▲	▲
Czech Republic	▲	▲	▲	▲	o	o	o
New Zealand	o	o	o	o	o	o	o
Australia	o	o	o	o	o	o	o
Canada	o	o	o	o	o	o	o
Austria	o	o	o	o	o	o	o
Denmark	▼	▼	▼	o	o	o	o
France	o	o	o	o	o	▲	o
Iceland	▼	▼	▼	▼	o	o	o
Germany	o	o	o	o	o	o	o
Sweden	o	o	o	o	▼	▼	▼
Poland	▲	▲	▲	▲	o	o	o
Latvia	▲	▲	▲	▲	▲	o	o
Norway	o	o	o	o	o	o	o
Hungary	o	o	o	o	o	o	▲
Spain	o	o	o	o	o	o	o
Ireland	**o**	**o**	**o**	**o**	**o**	**o**	**o**
Russian Federation	o	o	o	o	o	o	o
United States	o	o	o	o	o	o	o
Italy	o	o	o	▲	▲	▲	o
Portugal	▲	▲	▲	o	o	o	o
Greece	o	o	o	o	▼	▼	▼
Thailand	▲	▲	▲	▲	o	o	o
Mexico	o	▼	▼	▼	▼	o	o
Indonesia	▲	▲	▲	▲	o	o	o
Brazil	▲	▲	▲	▲	o	o	o
OECD Total	**o**	**o**	**o**	**o**	**o**	**o**	**o**
OECD Average	**o**	**o**	**o**	**o**	**o**	**o**	**o**

Note. OECD countries are in regular font, partner countries are in italics. Countries are ordered in descending order of mean score on the Space & Shape subscale. Comparisons take the equating error into account.

Comparisons are not possible for Luxembourg, the UK and the Netherlands.

Significantly higher in 2003 ($p \leq .05$) ▲

The same in 2000 and 2003 ($p > .05$) o

Significantly lower in 2003 ($p \leq .05$) ▼

Table 3.28. *Comparison of Mean Performance and Performance at Key Percentiles on the Change and Relationships Subscale, 2000 and 2003 – OECD and Partner Countries*

	Point on the scale being compared						
	5th	*10th*	*25th*	*Mean*	*75th*	*90th*	*95th*
Korea	o	o	o	▲	▲	▲	▲
Finland	o	o	o	▲	▲	▲	▲
Hong Kong - China	o	o	o	o	o	o	o
Liechtenstein	▲	▲	▲	▲	o	o	o
Canada	▲	▲	▲	▲	▲	▲	▲
Japan	o	o	o	o	o	o	o
Belgium	▲	o	o	▲	▲	▲	o
New Zealand	o	o	o	o	o	o	o
Australia	o	o	o	o	o	o	o
Switzerland	▲	▲	▲	o	o	o	o
France	o	o	o	o	o	o	o
Czech Republic	▲	▲	▲	▲	▲	▲	o
Iceland	o	o	o	o	o	o	o
Denmark	▲	o	o	o	o	o	o
Germany	▲	o	▲	▲	▲	▲	▲
Ireland	**o**	**o**	**o**	**o**	**o**	**o**	**o**
Sweden	o	o	o	o	o	▲	▲
Austria	o	o	o	o	o	o	o
Hungary	▲	▲	▲	▲	o	o	o
Norway	o	o	o	o	o	o	o
Latvia	▲	▲	▲	▲	o	o	o
United States	o	o	o	o	o	o	o
Poland	▲	▲	▲	▲	o	o	o
Spain	o	o	▲	▲	o	o	o
Russian Federation	▲	▲	▲	o	o	o	o
Portugal	o	o	o	▲	▲	▲	▲
Italy	o	o	o	o	o	▲	▲
Greece	▲	▲	o	o	o	o	▼
Thailand	▼	▼	▼	▼	o	▲	▲
Mexico	o	o	o	o	o	o	o
Indonesia	▼	▼	▼	o	o	▲	▲
Brazil	▲	▲	▲	▲	▲	▲	▲
OECD Total	o	o	o	o	o	o	o
OECD Average	▲	▲	o	▲	▲	▲	▲

Note. OECD countries are in regular font, partner countries are in italics. Countries are ordered in descending order of mean score on the change and relationships subscale. Comparisons take the equating error into account.

Comparisons are not possible for Luxembourg, the UK and the Netherlands.

Significantly higher in 2003 ($p \leq .05$) ▲
The same in 2000 and 2003 ($p > .05$) **o**
Significantly lower in 2003 ($p \leq .05$) ▼

Table 3.29. Comparison of Mean Performance and Performance at Key Percentiles on the Reading Literacy Scale, 2000 and 2003 – OECD and Partner Countries

	Point on the scale being compared						
	5th	*10th*	*25th*	*Mean*	*75th*	*90th*	*95th*
Finland	o	o	o	o	o	▼	▼
Korea	o	o	o	o	▲	▲	▲
Canada	o	o	o	o	▼	▼	▼
Australia	o	o	o	o	o	o	o
Liechtenstein	▲	▲	▲	▲	▲	▲	▲
New Zealand	o	o	o	o	o	o	o
Ireland	**o**	**o**	**o**	▼	▼	▼	▼
Sweden	o	o	o	o	o	o	o
Hong Kong-China	o	o	▼	▼	▼	▼	▼
Belgium	o	o	o	o	o	o	o
Norway	o	o	o	o	o	o	o
Switzerland	o	▲	o	o	o	o	o
Japan	▼	▼	▼	▼	o	o	o
Poland	▲	▲	▲	▲	o	o	o
France	▼	o	o	o	o	o	o
United States	o	o	o	o	o	o	▼
Denmark	o	o	o	o	▼	▼	▼
Iceland	▼	▼	▼	▼	▼	o	o
Germany	o	o	o	o	o	o	o
Austria	▼	▼	▼	▼	o	o	o
Latvia	▲	▲	▲	▲	▲	▲	o
Czech Republic	o	o	o	o	o	o	o
Hungary	o	o	o	o	o	o	o
Spain	▼	▼	▼	▼	o	o	o
Portugal	o	o	o	o	o	o	o
Italy	▼	▼	▼	▼	o	o	o
Greece	o	o	o	o	o	o	o
Russian Federation	▼	▼	▼	▼	▼	▼	▼
Thailand	o	o	▼	o	o	o	o
Brazil	▼	▼	o	o	▲	▲	▲
Mexico	▼	▼	▼	▼	▼	o	o
Indonesia	o	o	o	o	o	o	o
OECD Total	▼	▼	▼	▼	**o**	**o**	**o**
OECD Average	**o**	**o**	**o**	**o**	**o**	**o**	**o**

Note. OECD countries are in regular font, partner countries are in italics. Countries are ordered in descending order of mean score on the reading literacy scale. Comparisons take the equating error into account.

Comparisons are not possible for Luxembourg, the UK and the Netherlands.

Significantly higher in 2003 ($p \leq .05$) ▲

The same in 2000 and 2003 ($p > .05$) **o**

Significantly lower in 2003 ($p \leq .05$) ▼

Table 3.30. Comparison of Mean Performance and Performance at Key Percentiles on the Science Scale, 2000 and 2003 – OECD and Partner Countries

	Point on the scale being compared						
	5th	*10th*	*25th*	*Mean*	*75th*	*90th*	*95th*
Finland	o	o	o	▲	▲	▲	▲
Japan	▼	▼	▼	o	o	▲	▲
Hong Kong-China	o	o	o	o	o	o	o
Korea	▼	▼	▼	▼	o	o	▲
Liechtenstein	o	o	▲	▲	▲	▲	▲
Australia	▼	o	o	o	o	o	o
Mexico	▼	▼	▼	▼	o	o	o
Czech Republic	o	o	o	▲	▲	▲	▲
Canada	▼	▼	▼	▼	o	o	o
Switzerland	o	o	o	▲	▲	▲	▲
France	o	o	o	▲	▲	▲	▲
Belgium	o	o	o	▲	▲	▲	▲
Sweden	▼	▼	o	o	o	▲	o
Ireland	**o**	**o**	**o**	**o**	**o**	**o**	**o**
Hungary	o	▲	o	o	o	o	o
Germany	o	o	o	▲	▲	▲	▲
Poland	o	o	o	▲	▲	▲	▲
Iceland	▼	o	o	o	o	o	o
United States	o	o	o	o	o	o	o
Austria	▼	▼	▼	▼	▼	▼	▼
Russian Federation	▲	▲	▲	▲	▲	▲	▲
Latvia	▲	▲	▲	▲	▲	▲	▲
Spain	o	o	o	o	o	o	o
Italy	o	o	o	o	▲	▲	▲
Norway	▼	▼	▼	▼	▼	o	o
Greece	o	o	▲	▲	▲	▲	▲
Denmark	o	o	o	o	o	o	o
Portugal	o	o	o	o	o	o	o
Thailand	o	▼	▼	o	o	o	o
New Zealand	o	o	o	o	o	o	o
Indonesia	o	o	o	o	o	o	o
Brazil	o	o	o	▲	▲	▲	▲
OECD Total	▼	▼	▼	**o**	**o**	**o**	▲
OECD Average	▼	▼	**o**	**o**	**o**	▲	▲

Note. OECD countries are in regular font, partner countries are in italics. Countries are ordered in descending order of mean score on the science scale. Comparisons take the equating error into account.

Comparisons are not possible for Luxembourg, the UK and the Netherlands.

Significantly higher in 2003 ($p \leq .05$) ▲

The same in 2000 and 2003 ($p > .05$) **o**

Significantly lower in 2003 ($p \leq .05$) ▼

Performance on the Science Scale in 2000 and 2003

Table 3.30 compares the mean performance in 2000 and 2003 of countries in science and also compares performance at a number of percentile points.

The OECD country average score and OECD average scores at the 25th and 75th percentiles did not change. However, a significant decrease at the lower end of the distribution and a significant increase at the upper end are in evidence, indicating a wider range of achievement in 2003 than in 2000. There is no change at any point of the distribution for Ireland.

RELATING PERFORMANCES IN THE THREE ASSESSMENT DOMAINS

The achievements of students in combined mathematics, reading literacy, and science in PISA 2003 are highly interrelated (Table 3.31). Unadjusted correlations are at or above .80 for each comparison (.80 mathematics and reading literacy; .84 for mathematics and science; and .85 for reading literacy and science). Inter-correlations among the mathematics subscales are marginally higher, ranging from .87 to .93. Correlations among the three domains in PISA were similarly high in PISA 2000, ranging from .82 (reading literacy and mathematics) to .90 (reading literacy and science). Correlations between these domains and cross-curricular problem solving in PISA 2003 are examined in Chapter 7.

Table 3.31. Associations Between the Combined Mathematics, Reading Literacy, and Science Scales, and the Four Mathematics Subscales (Irish Students)

	Raw Coeff	*SD Unit*	*SE*	*r*	*t*	*p*
Combined Scales						
Comb. Mathematics-Reading	0.81	69.05	0.02	**.798**	52.36	<.001
Comb. Mathematics-Science	0.77	65.60	0.01	**.840**	71.61	<.001
Reading-Science	0.79	69.00	0.01	**.847**	63.46	<.001
Mathematics Subscales						
Space & Shape-Change & Rel.	0.94	88.93	0.01	**.872**	82.08	<.001
Space & Shape-Uncertainty	0.93	88.31	0.01	**.879**	76.92	<.001
Space & Shape-Quantity	0.94	88.75	0.01	**.877**	76.42	<.001
Change & Rel.-Uncertainty	0.91	79.96	0.01	**.928**	115.29	<.001
Change & Rel.-Quantity	0.91	80.04	0.01	**.922**	104.26	<.001
Uncertainty-Quantity	0.91	81.93	0.00	**.913**	281.98	<.001

Note. The column "SD Unit" shows the increase in the outcome variable per standard deviation of the increase in the explanatory variable.

Standard deviations for scales are as follows: combined mathematics 85.3, reading literacy 87.5, science 93.0, Space and Shape 94.5, Change and Relationships 87.5, Quantity 88.2, Uncertainty 89.8.

CONCLUSION

In this chapter, the achievement outcomes of students in PISA 2003 in mathematics, reading literacy, and science were described, with a particular focus on the achievements of Irish students. Mean performance and the distribution of performance in the three subject domains were examined. Achievements on four mathematics subscales were also described. The performance of students in 2000 and 2003 on two aspects of mathematics, on reading and on science was compared.

The mean performance of Irish students on the combined mathematics scale in 2003 did not differ significantly from OECD average performance. Ireland ranked 17th out of 29 OECD countries (95% confidence interval of Ireland's ranking = 15th to 18th) and 20th out of all 40 participating countries that met sampling requirements (95% confidence interval of Ireland's ranking =17th to 21st). Taking sampling and measurement error into account, Ireland's mean score was significantly lower than that of 10 OECD countries (or 13 of all participating countries) and significantly higher than that of 10 OECD countries (or 18 of all participating countries). Using a nonparametric maximum likelihood estimation method which arranges countries into empirically distinct performance groupings, Ireland's performance was indistinguishable from that of France, Sweden, Austria, Germany, and the Slovak Republic, falling in the second highest of six groups.

Although some countries showed consistently high performance on all four mathematics subscales (e.g., Finland, Korea, the Netherlands), Ireland's performance showed some variation. Irish students achieved a mean of 476.2 points on the Space & Shape subscale – a score that was significantly below the corresponding OECD country average score (496.3). In contrast, Ireland's mean score on the Change & Relationships subscale (506.0), although just about seven points (about one-twentieth of a standard deviation) above the OECD average, was nonetheless significantly higher than it. Performance on the Quantity subscale (501.7) did not differ significantly from the OECD average (500.7). Irish students performed best on the Uncertainty subscale, achieving a mean score of 517.2 (about 15 points, and significantly above, the OECD average of 502.0). Comparisons of the mean performance of countries using the nonparametric maximum likelihood estimation method showed that, across the four subscales, country level performance was negatively skewed, with the suggestion of a long tail at the lower end of the distribution of masspoints (mean scores for country groupings).

Six categorical proficiency levels were developed for the combined mathematics scale and the four subscales. These permit a description of the skills associated with representative test items at each level. Upper and lower ends were unbounded, meaning that PISA does not assess all the likely skills of students who are at the upper end of Level 6 (the highest level), or the skills of students who did not demonstrate sufficient proficiency to meet the requirements of Level 1.

An examination of the percentage of Irish students who scored at each proficiency level with reference to the OECD average suggests that Irish performance was characterised by comparatively fewer low achievers and comparatively fewer high achievers. For example, 16.8% of Irish students were at or below Level 1 on the combined mathematics scale, compared to just over 21% of students on average across the OECD. In contrast, just over 11% of Irish students scored at Levels 5 or 6, compared to an OECD average of 14.6%, despite the fact that the Irish mean performance does not differ from the OECD average mean score. This contrasts with the performance of some other countries: for example, although the mean scores for Germany and Ireland are similar, there are comparatively more students in Germany at or below Level 1 (21.6%) and at Levels 5 and 6 (16.3%).

The high percentage of Irish students at or below Level 1 on the Space & Shape subscale (27.6% compared to an OECD average of 24.8%), as well as the low percentage scoring at Levels 5 and 6 (8.6% compared to an OECD average of 16.2%) further demonstrate the comparatively poor performance of Irish students on this subscale. Irish students fared somewhat better on the Change & Relationships subscale (with 16.3% at or below Level 1 and almost 13% at Levels 5 and 6) as well as on the Uncertainty subscale (with 13.8% at or below Level 1 and just over 16% at Levels 5 and 6). On the Quantity subscale, 17.9% of Irish students were at Level 1 or below (compared to an OECD average of 21.3%). The combined percentage at Levels 5 and 6 (just over 16%) is close to the OECD average of 15.0%.

The standard deviation, which gives an indication of the dispersion of scores, was smaller for Ireland when compared with the OECD average on the combined mathematics scale and on all four subscales. For example, the Irish standard deviation for the combined scale is 85.3 compared to the OECD average of 100.0 (range of standard deviations for all countries = 80.5 to 109.9). This comparatively narrow dispersion of scores in Ireland is also evident in the difference between scores at the 10th and 90th percentiles. The OECD average difference between these two points for the combined scale is 259.3 points; for Ireland, it is 220.8 points, indicating a higher level of similarity in achievement outcomes between students in Ireland. Again, the same pattern is evident in the case of all four mathematics subscales.

On the combined mathematics scale, the score at the 10th percentile in Ireland was 24.1 points (about three-tenths of a standard deviation) higher than the OECD average (393.1 compared to 369.0); at the 90th percentile it was 14.4 points (about one-sixth of a standard deviation) lower (613.9 compared to 628.3). This again demonstrates a relatively strong performance at the lower end of the distribution, and the relatively weak performance of students at the upper end. The performance of higher-performing Irish students was similar to that of the OECD average at the 90th percentile only in the case of the Uncertainty subscale.

Ireland’s mean score (515.5) on reading literacy in 2003 was significantly higher than the OECD average of 494.2, giving it a ranking of 6th of 29 OECD countries (95% confidence interval of Ireland’s ranking = 6th to 8th), and 7th out of 40 countries (95% confidence interval of Ireland’s ranking = 6th to 10th). Just three countries (Finland, Korea, and Canada) scored significantly higher. In analyses using the nonparametric maximum likelihood estimation method, 10 distinct performance groupings emerged. Ireland’s mean score fell into the third grouping, and is indistinguishable from those of Sweden, the Netherlands, Hong Kong-China, and Belgium.

When performance on the reading literacy scale is considered in terms of categorical proficiency levels similar to those developed for mathematics (but with five levels rather than six), the proportion of Irish students scoring at or below Level 1 (11.0%) is lower than the corresponding OECD country average (19.1%). More Irish students score at Level 5 (9.3%) compared to the corresponding OECD average (8.3%). Ireland’s score at the 10th percentile is 40.5 points higher than the OECD average score at this point (401.3 compared to just 360.8), but at the 90th percentile it is just 5.2 points higher (622.1 compared to 616.9). While the OECD average score difference between these two percentile points is 256.1, the difference for Ireland is comparatively low at just 220.8 points, indicating a relatively narrow spread in reading achievement. As with mathematics, the difference between the highest and lowest scoring countries (about 170 points) is much smaller than the differences between the OECD average scores at the 10th and 90th percentiles, indicating greater variation in reading literacy within countries than between countries.

Ireland's mean score of 505.4 for science in 2003, although just about 5 score points above the OECD average of 499.6, is nonetheless significantly higher. Ireland ranks 13th of 29 OECD countries (95% confidence interval of Ireland's ranking = 9th to 16th) and 16th of 40 countries (95% confidence interval of Ireland's ranking = 12th to 30th). When country mean scores for science are grouped into performance groupings using nonparametric maximum likelihood estimation, eight distinct groups emerge. Ireland is in the third highest grouping, and its mean performance is indistinguishable from that of Belgium, Sweden, Hungary, and Germany. An examination of scores at the 10th and 90th percentiles reveals a pattern similar to that for combined mathematics. The score at the 10th percentile for Ireland (383.9) is 22.3 points higher than the OECD country average score at this point (361.6), while the Irish score at the 90th percentile (624.5) is just 9.7 points lower than the OECD country average (634.2). Again, there is a comparatively narrow spread of scores between the 10th and 90th percentiles in Ireland (240.6 compared to an OECD average of 272.6).

A comparison between student performance in 2000 and 2003 on the Space & Shape subscale revealed no change in OECD country average performance or in the performance of Irish students in terms of mean scores, or in terms of scores at key percentile ranks. On the other hand, although there were no changes in Ireland in the Change & Relationships scale, there was an overall increase in OECD country average performance, as well as increases in performance at five percentile points.

The OECD country average score, and the OECD average scores at key percentile points remained unchanged in reading literacy between 2000 and 2003. However, the mean score for Ireland was significantly lower in 2003 than in 2000 (by 11.2 score points). The performance of Irish students also decreased significantly at the 75th, 90th, and 95th percentile points. Further data, in PISA 2006 and beyond, will be required to make any strong inferences regarding a possible trend.

Although there was no change in the OECD country average score in science between 2000 and 2003, a wider spread in achievement is in evidence in 2003, with lower scores at the 5th, 10th and 25th percentiles and higher scores at the 75th, 90th and 95th percentiles. The mean score for Ireland for science did not change in 2003, nor did scores at any of these percentile points.

4

Student- and School-Level Associations with Achievement in Combined Mathematics, Reading and Science in PISA 2003

This chapter examines associations between of a range of explanatory variables and student achievement in three PISA 2003 assessment domains – combined mathematics, reading and science. The explanatory variables include student-level characteristics such as gender and motivation and school-level characteristics such as school socioeconomic composition, climate, and resources (Inset 4.1).

The chapter is divided into four sections. In the first, achievement outcomes associated with various student characteristics (e.g., gender) are reported and compared to one another. In the second section, analyses are reported for school-level variables. In the third section, correlations of student and school variables with achievement are considered. The fourth and final section examines associations among the explanatory variables themselves, at both the student and school levels. While the primary focus of the chapter is on describing relationships between explanatory variables and the achievement of Irish students in PISA 2003, reference is made at appropriate points to relationships for students in other countries.

The analyses reported in this chapter examine associations between pairs of variables – a single explanatory variable and a response variable (for example, between gender and achievement in mathematics). In Chapter 5, the results of more complex multilevel analyses are presented. These seek to explain variation in achievement between and within schools by examining the simultaneous impact of a number of student and school variables on achievement scores. The particular variables selected for the analyses in this chapter and in Chapter 5 are based on a review of previous research, on the priorities for analysis identified by the Irish PISA National Advisory Committee, and the proportion of data available on individual items (i.e., low levels of missing values).

Explanatory variables may be classified according to whether they are categorical or continuous. Categorical variables typically have two or more discrete categories (for example, male/female; small/medium/large). Continuous variables describe a quantity (for example, the time spent on mathematics homework per week). Some continuous variables are composites based on two or more discrete variables. The composites were formed by first conducting a principal components analysis among several items, and then applying a one-parameter (Rasch) Item Response Theory model to generate weighted likelihood estimates (scores) on a scale with a mean of zero and a standard deviation of 1 across all OECD countries. Examples of composite variables that were computed in this manner are the index of economic, social and cultural status, and self-reported anxiety about mathematics. A description of questionnaire items associated with each explanatory variable is provided in Appendix B (Section B.1). The reader is referred to Insets 4.2 and 4.3 for a discussion of some of the technical issues that arise in performing and interpreting the analyses reported in this chapter.

OVERVIEW OF STUDENT- AND SCHOOL-LEVEL VARIABLES

The range of student and school characteristics considered in this chapter are detailed in Inset 4.1. At the student level, for the most part, data were gathered/derived from the Student Questionnaire, which was administered to all students who participated in the assessment. At the school level, data were gathered/derived from the School Questionnaire, completed by principal teachers of participating schools, as well as from the post-primary school database of the Department of Education and Science and the Junior Certificate Examination database of the State Examinations Commission.

Four broad categories of student-level characteristics are considered: student demographics/home background, out-of-school activities, academic characteristics and behaviour, and attitudes towards mathematics. At the school level, two broad categories of characteristics are considered: school structure and composition, and school climate and resources. Additional variables are considered in the OECD reports on PISA 2003 (OECD, 2004b, c), and the full range of variables gathered/derived in the course of the survey can be found in the PISA 2003 database.[22] In considering variables derived from the School Questionnaire, it should be borne in mind that just 133 out of 145 School Questionnaires distributed to schools in Ireland were returned. Although this represents a return rate of 91.7%, it does result in a rate of missing data at the student level of 10% or more for all variables drawn from the School Questionnaire. Four variables reported at the school level (school economic, social and cultural status; perceived disciplinary climate; total instructional time; and instructional time for mathematics) were originally collected at the student level. In these four instances, the data were aggregated to the school level and disaggregated back to the student level (i.e., each student was assigned the mean for his or her school).

Inset 4.1. Student and School Characteristics

Student Characteristics

Demographics/Home Background
- Gender
- Nationality
- Membership of the Traveller Community [N]
- Parents' Occupation (Socioeconomic Status)
- Parents' Education
- Economic, Social and Cultural Status
- Family Structure
  - Household Composition
  - Number of Siblings [N]
- Home Educational Resources [N]
- Number of Books in the Home

Out-of-School Activities
- Homework Practices
  - Total Time on Homework/Study
  - Total Time on Mathematics Homework/Study
- Leisure Reading
  - Frequency of Reading Fiction [D]
  - Frequency of Reading E-mails/Webpages [D]

Attitudes Towards Mathematics
- Self-efficacy in Mathematics [D]
- Anxiety about Mathematics [D]

Student Characteristics (continued)

Academic Characteristics/Behaviour
- Current Grade Level
- Absence from School [N]
- Risk of Early School Leaving [N]
- Syllabus Level Taken at Junior Cert [D, N]
- Use of Calculator in PISA[D]
- Study of Science [D]

School Characteristics

Structure and Composition
- Stratum (Size)
- Sector [N]
- Disadvantaged Status [N]
- Gender Composition [N]
- Economic, Social and Cultural Status
- Fee Waiver Entitlement [N]

Climate and Resources
- Disciplinary Climate – Mathematics [D]
- Ratio of Computers to Students
- Total Instructional Time
- Instructional Time in Mathematics [D]

[D] Domain-specific variable, [N] National variable

[22] This is available, together with supporting documentation, at http://www.pisa.oecd.org.

Inset 4.2. A Note on the Analyses

Weighting of responses. All percentages and mean achievement scores reported in this chapter are estimates that were computed using normalised population weights. The standard errors accompanying mean achievement scores were computed using a balanced repeated replication (BRR) method of variance estimation that took the clustered nature of the PISA sample design into account.

Categorisation of continuous variables. For descriptive purposes, continuous variables (including weighted likelihood estimates) were divided into high, medium, and low using their 33rd and 67th percentiles as cut-points. In some cases the percentage of students represented in a category does not correspond exactly to one-third of available cases because of tied ranks at the designated cut-points or due to the discrete nature of a particular variable.

Treatment of missing values. Two types of percentages are reported for each variable for which there are missing cases. The first of these (%T – percentages of all students), relates to all students for whom achievement data are available, including those for whom data on the questionnaire variable are missing. The second (%A – percentages of available students) refers to students for whom achievement data and data on the variable are available, and does not include missing cases. Where percentages are discussed in the text, they refer to the first column (i.e., to total cases). For most variables, the percentage of missing values is less than five. Where missingness for a variable exceeds 5%, the 'missing' category is discussed in the text, and reference is made to the performance of students in the missing category relative to the reference category.

Testing for the statistical significance of mean score differences. Tests designed to ascertain the statistical significance of differences between mean achievement scores associated with different levels of each explanatory variable were conducted. The statistical procedure used to determine whether a difference between groups is significant required the selection of an appropriate reference category for each variable. Comparisons were then made between the mean score of the reference category and the mean score of each remaining category, including, where relevant, the missing value category.

To reduce the possibility of making a Type 1 (*alpha*) error (i.e., incorrectly inferring a significant difference) in the context of conducting multiple comparisons, Bonferroni's procedure was applied and appropriately adjusted critical (*t*) values corresponding to the .05 and .10 levels were obtained in WesVar (see Section A4.4, Appendix 4). Then, confidence intervals were constructed by adding to and subtracting from each mean score the difference of the product of the corresponding standard error of the difference and the adjusted critical value. Although not reported in the tables in this chapter, 90% confidence intervals were constructed to identify any differences which, though not significant at the conventional .05 level, might be significant at the .10 level. Such differences are reported in the text since they may be worthy of exploration in future research. It can be concluded that a difference is statistically significant if the relevant confidence interval does not include the value of zero.

Testing for the statistical significance of the difference between percentages. A similar approach to that used to test the significance of mean score differences was used to test the significance of the difference between pairs of percentages (for example, the percentages of male and female students at a particular proficiency level on the combined mathematics scale). The large-sample Normal sampling distribution (rather than the *t* distribution) was used to obtain the 90% and 95% confidence intervals, as this avoids the complexities involved in calculating the degrees of freedom corresponding to values of *t* (Agresti & Finlay, 1997, pp. 219-222).

STUDENT CHARACTERISTICS AND ACHIEVEMENT

Student Gender

Gender Differences in Mean Scores

Almost equal percentages of Irish male and female students were assessed in mathematics, reading and science in PISA 2003 (Table 4.1; see also Inset 4.3). Male students achieved a mean score in mathematics that was one-sixth of a standard deviation higher than that of females.[23] In contrast, females achieved a mean score in reading literacy that was one-third of a standard deviation higher than the mean score of males. Both of these differences are statistically significant. The difference between the mean scores of males and females in science (a mere 2 scale score points) is not statistically significant.

Male students achieved significantly higher mean scores than their female counterparts in 21 of 29 OECD countries (and 27 of 40 participating countries) on the combined mathematics scale, with an average OECD country difference of 11.1 points (Appendix B, Table B.1). Iceland is the only country in which female students had a significantly higher mean score than males (by 15.4 points).

On the reading scale, female students achieved significantly higher mean scores than males in all countries, with the exception of partner country Liechtenstein. The OECD average difference of 34.1 points favouring females was larger than the OECD average country difference favouring males for mathematics. The difference in reading of 29.0 points (three-tenths of a standard deviation) in Ireland is similar to the OECD country average difference. Overall significant differences ranged from 57.7 points in Iceland to 13.3 points in partner country Macao-China.

For science, no clear pattern of gender differences emerged across countries, with an OECD average of just 5.8 points favouring female students. In four countries, the gender difference in mean science achievement was zero. Overall, significant differences favouring males were reported for Iceland and Tunisia only, while significant differences favouring females were reported for 12 countries, most notably partner country Liechtenstein, where female students outperformed their male counterparts by 26.0 points (Table B.1, Appendix B).

Table 4.1. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Gender

	Frequencies (All Domains)	*Maths (Comb.)*		*Reading*		*Science*	
	%T	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Female	49.7	495.4	3.39	530.1	3.71	504.4	3.88
Male	50.3	510.2	3.01	501.1	3.26	506.4	3.08
All Available	100.0	504.8	2.41	517.4	2.55	502.9	2.65

Mean Score Differences (Reference Category: Male)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Fem–Male	-14.8	4.19	**-23.1**	**-6.5**	29.0	4.56	**19.9**	**38.1**	-2.0	4.48	-10.9	7.0

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

[23] Where differences are reported in terms of standard deviation units, these refer to the Irish standard deviations, which vary by domain and subscale (see Tables 3.1, 3.19 and 3.24 in Chapter 3).

Inset 4.3. Identifying a Significant Difference between Mean Achievement Scores

Throughout this chapter, reference is made to differences between mean achievement scores. As indicated in Inset 4.2, the approach taken to examining whether or not a difference between mean scores is statistically significant involved computing the standard error of the difference, identifying the relevant critical values (*t* scores) adjusted for multiple comparisons where appropriate, and constructing 95% and 90% confidence intervals around the difference. An example of how differences between mean scores may be interpreted is provided here.

Mean Score Differences (Reference Category: Male)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Fem-Male	-14.8	4.19	**-23.1**	**-6.5**	29.0	4.56	**19.9**	**38.1**	-2.0	4.48	-10.9	7.0

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

In this example (from Table 4.1), *critical* alpha values corresponding to the 95% and 90% confidence intervals (i.e., .05 and .10) were obtained from a table of critical values of *t*, using an approximation of 80 degrees of freedom (the number of variance strata associated with balanced repeated replication (BRR) method of variance estimation). The obtained critical value (1.990) was then multiplied by the standard error of the difference (computed in WesVar 4.2) (4.19), and added to and subtracted from the mean score difference. For combined mathematics in our example, the 95% confidence interval is –23.1 to –6.5. Since 0 (i.e., no difference) is outside this interval, it can be concluded with 95% confidence that the obtained difference of –14.8 between males and females is statistically significant. The confidence intervals corresponding to such differences are indicated in bold. In the case of science, where the difference between females and males is not statistically significant at the .05 level, a 90% confidence interval was also constructed (–9.4 to 5.5). Again, the difference of 2.0 (in favour of males) is not statistically significant as 0 falls within the obtained interval.

In some tables there is more than one comparison for each domain. Where this arises, the *alpha* level is adjusted by dividing it by the number of comparisons to be made, and locating the corresponding critical *t* value in a table of critical values of *t* (again using 80 degrees of freedom). In the case of Table 4.7, for example, where there are two comparisons for each domain, alpha (.05) was divided by two, resulting in an adjusted *alpha* of .025. Most tables, including Table 4.7, contain a ‘missing’ category, denoting the proportion of respondents for whom questionnaire data on the item under consideration were not available. In general, the mean score of students in this category is compared to the reference category to ascertain if the missing group differ in achievement from the reference category.

The performance of male and female students on the four mathematics subscales was also examined. Irish male students significantly outperformed females on all four subscales. The gender difference in Ireland is greatest for the Space & Shape subscale and smallest for the Quantity subscale. Male students significantly outperformed female students, by just over one-quarter of an Irish standard deviation (25.5 points) on the Space & Shape subscale, by one-sixth of a standard deviation on both the Change & Relationships and the Uncertainty subscales (12.6 and 15.5 points, respectively), and by about one-tenth of a standard deviation (8.9 points) on the Quantity scale (Table 4.2a, 4.2b).

Across all countries, gender differences tend to be largest on the Space & Shape subscale, with an OECD country average difference of 16.7 points, but, among OECD countries, the Czech Republic, Korea, Luxembourg, and Switzerland had differences on the subscale that were equivalent to or greater than in Ireland. The OECD country average difference for Change & Relationships is 11.0 points, for Quantity, it is just 6.2 points, and for Uncertainty, it is 12.6 points (Table 4.2a, b). Although statistically significant for each scale, OECD average differences are small, and mask large differences between males and females within some countries. For example, male students in Korea achieved mean scores on all four subscales that were between 22 to 28 points higher than those of female students in the same country (see Table B.2, Appendix B).

Table 4.2a. Mean Scores of Irish Students for Space and Shape, and Change and Relationships, and Mean Score Differences, by Gender

	Frequencies (All Subscales)	*Space/Shape*		*Change/Rel.*	
	%T	*Mean*	*SE*	*Mean*	*SE*
Female	49.6	463.4	3.44	499.6	3.52
Male	50.4	488.9	2.96	512.2	3.05
All Available	100.0	476.2	2.43	506.0	2.45

Mean Score Differences (Reference Category: Male)

	Space/Shape				*Change/Relationships*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Fem–Male	-25.5	4.28	**-34.0**	**-16.9**	-12.6	4.43	**-21.5**	**-3.8**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Table 4.2b. Mean Scores of Irish Students for Quantity, and Uncertainty, and Mean Score Differences, by Gender

	Frequencies (All Subscales)	*Quantity*		*Uncertainty*	
	%T	*Mean*	*SE*	*Mean*	*SE*
Female	49.6	497.2	3.47	509.4	3.73
Male	50.4	506.1	3.06	524.9	3.24
All Available	100.0	501.7	2.48	517.2	2.65

Mean Score Differences (Reference Category: Male)

	Quantity				*Uncertainty*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Fem–Male	-8.9	4.28	**-17.4**	**-0.4**	-15.5	4.60	**-24.6**	**-6.3**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Gender Differences on the Mathematics Proficiency Levels

Gender differences in mathematics were also examined in terms of the proportions of males and females at each proficiency level on the combined mathematics scale and on each of the four mathematics subscales. As indicated in Chapter 3 (Inset 3.3), each proficiency level encompasses a range of skills (represented by items at that level), and students scoring at a particular level have an average estimated probability of .62 of succeeding on items at that level. The mathematical knowledge of students scoring below Level 1 is not assessed by PISA.

Although male students in Ireland achieved a significantly higher mean score on the combined mathematics scale, this difference is not apparent in the proportions of males and females achieving each proficiency level on the combined mathematics scale (see Table 4.3). Hence, although there are fewer Irish male students (4.2%) than females (5.2%) with scores that are below Level 1, the difference (about 1%) is not statistically significant. Similarly, although there are more females (24.7%) than males (22.5%) at Level 2, the difference is not significant. At Level 5, there is a borderline statistically significant difference between the percentage of males (10.8%) and females (7.4%) (Bonferroni 90% Confidence Interval: 0.04 to 6.8).

Table 4.3. Percentages of Irish Students at Each Combined Mathematical Proficiency Level and Percentage Differences, by Gender

	Males		*Females*		*All*	
Level	*Percent*	*SE*	*Percent*	*SE*	*Percent*	*SE*
Below Level 1	4.2	0.79	5.2	0.74	4.7	0.57
Level 1	10.8	1.13	13.5	1.28	12.1	0.84
Level 2	22.5	1.44	24.7	1.37	23.6	0.83
Level 3	27.8	1.46	28.2	1.36	28	0.82
Level 4	21.0	1.63	19.4	1.21	20.2	1.06
Level 5	10.8	1.09	7.4	0.83	9.1	0.76
Level 6	2.9	0.5	1.6	0.36	2.2	0.33
Total	100.0		100.0		100.0	

Percentage Difference (Reference Category: Male)

	Difference	*SED*	*BCI95%*	
Below Level 1	1.0	1.08	-1.9	3.9
Level 1	2.7	1.71	-1.9	7.3
Level 2	2.2	1.99	-3.1	7.5
Level 3	0.4	2.00	-5.0	5.8
Level 4	-1.6	2.03	-7.1	3.9
Level 5	-3.4	1.37	-7.1	0.3
Level 6	-1.3	0.62	-3.0	0.4

Note. SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted confidence intervals

There were, however, some significant differences between the proportions of Irish male and female students at various proficiency levels on the mathematics subscales (Appendix B, Tables B.3 to B.6). Consistent with overall mean gender differences, the most notable differences were found for the Space & Shape subscale, where significantly more females (13.0%) than males (8.6%) performed below Level 1, and significantly more males (2.5%) than females (1.1%) performed at Level 6 (Table B.3). Significant gender differences favouring males were also associated with Level 6 on the Uncertainty scale (Table B.6).

Gender Differences on the Reading Proficiency Levels

A comparison of the proportions of males and females at each reading proficiency level indicated that significantly more females than males achieved Levels 4 and 5 (the highest levels), while significantly more males than females achieved Levels 1 and 2 (Table 4.4). Although there were more males (3.6%) than females (1.8%) below Level 1, the difference was not statistically significant.

Table 4.4. Percentages of Irish Students at Each Reading Literacy Proficiency Level and Percentage Differences, by Gender

	Males		*Females*		*All*	
Level	*Percent*	*SE*	*Percent*	*SE*	*Percent*	*SE*
Below Level 1	3.6	0.68	1.8	0.51	2.7	0.48
Level 1	10.7	1.07	5.9	0.75	8.3	0.66
Level 2	24.1	1.43	18.2	1.44	21.2	1.2
Level 3	32.4	1.67	32.3	1.6	32.4	1.26
Level 4	22.9	1.57	29.5	1.59	26.2	1.19
Level 5	6.3	0.8	12.3	1.13	9.3	0.71
Total	100.0		100.0		100.0	

Percentage Difference (Reference Category: Male)

	Difference	*SED*	*BCI95%*	
Below Level 1	-1.8	0.85	-4.1	0.5
Level 1	-4.9	1.31	**-8.4**	**-1.4**
Level 2	-5.9	2.03	**-11.3**	**-0.4**
Level 3	-0.1	2.31	-6.3	6.2
Level 4	6.6	2.24	**0.5**	**12.6**
Level 5	6	1.39	**2.3**	**9.7**

Note. SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Gender Differences at Key Intervals of the Science Scale

Unlike mathematics and reading literacy, it was not possible to describe the performance of male and female students on proficiency levels in science, as levels have not yet been developed for this domain. However, it was possible to compare the proportions of male and female students achieving scores at different intervals along the science scale (e.g., the proportions scoring below the 10th percentile). Consistent with the finding that there is no significant gender difference associated with the overall mean score, there are no significant differences in the proportions of males and females at any of these scale intervals (Table 4.5).

Table 4.5. Percentages of Irish Students within Key Percentile Intervals in Science, and Percentage Differences, by Gender

	Males		*Females*		*All*	
Scale Interval	*Percent*	*SE*	*Percent*	*SE*	*Percent*	*SE*
<10th	10.0	1.07	10.0	1.05	10.0	0.49
10th-25th	15.3	1.76	14.8	1.74	15.0	0.48
26th-50th	24.5	1.63	25.4	1.83	25.0	0.58
51st-75th	24.5	1.28	25.7	1.32	25.1	0.59
76th-90th	15.1	1.02	14.8	1.31	15.0	0.44
>90th	10.6	0.8	9.3	0.94	9.9	0.38
Total	100.0		100.0		100.0	

Percentage Difference (Reference Category: Male)

	Difference	*SED*	*BCI95%*	
<10th	-1.8	-0.05	1.5	-4.0
10th–25th	-4.9	0.53	2.5	-6.0
26th–50th	-5.9	-0.83	2.5	-7.3
51st–75th	-0.1	-1.19	1.8	-6.0
76th–90th	6.6	0.28	1.7	-4.1
>90th	6.0	1.26	1.2	-2.0

Note. SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted confidence intervals

Comparison of Gender Differences in PISA 2000 and PISA 2003

In PISA 2000, a significant gender difference in favour of females in Ireland (28.7 points, or three-tenths of a standard deviation) was found for reading literacy, while Irish males significantly outperformed females in mathematics (by 12.9 points, or one-sixth of a standard deviation). There was no significant gender difference associated with performance on science in PISA 2000 (the difference, 6.2 scale points, was less than one-tenth of a standard deviation). The data indicate that the pattern of gender differences in overall performance in Ireland is similar for both PISA 2000 and PISA 2003.

Country of Birth

In response to a question asked of students in all participating countries in PISA 2003, 95.0% of students in Ireland indicated that they were 'native' students (i.e. they and at least one of their parents had been born in Ireland). Just under 1% indicated that they were 'first-generation' students (i.e., they were born in Ireland, but their parent(s) was/were foreign-born), while 2.5% described themselves as 'non-native' (i.e., the students and their parent(s), were foreign-born) (Table 4.6). No significant differences in achievement on combined mathematics, reading or science were observed between first-generation and native students, or between non-native and native students. In other OECD countries in which there are larger percentages of first-generation and non-native students, large differences in achievement between native on the one hand, and first generation and non-native students on the other, were observed (see OECD, 2004b, Table 4.2f). In Sweden for example, where 5.7% of students are classified as 'first-generation', native students outperformed first-generation students in combined mathematics by 34.5 points, while in France, where 10.8% of students are categorised as 'first-generation', the difference in favour of native students for combined mathematics was 47.8 points.

Table 4.6. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Native Status

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Native	95.0	96.5	503.2	2.44	516.0	2.62	505.6	2.66
First-Generation	0.9	1.0	474.4	19.17	486.7	22.14	488.1	23.18
Non-Native	2.5	2.5	508.7	11.79	511.2	12.01	504.6	13.77
Missing	1.6	0.0	488.8	25.88	510.5	31.26	502.7	30.26
All Available	98.4	100.0	506.9	2.46	519.9	2.58	505.1	2.69

Mean Score Differences (Reference Category: Native)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
First-Gen–Native	-28.79	19.20	-75.716	18.1	-29.29	22.03	-83.14	24.6	-17.54	22.98	-73.72	38.6
Non-Nat.–Native	5.5	11.89	-23.6	34.5	-4.7	12.12	-34.4	24.9	-1.0	13.92	-35.0	33.0
Missing–Native	-14.4	25.79	-77.5	48.6	-5.4	31.32	-82.0	71.1	-2.9	30.22	-76.8	70.9

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Membership of the Traveller Community

Less than one percent of students in Ireland indicated that they were members of the Traveller community (Table 4.7). Across each of the three assessment domains, these students performed significantly less well than students in the settled community. Differences ranged from three-quarters of a standard deviation (63.6 points) in mathematics to just over one standard deviation (85.4 points) in reading.

Table 4.7. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Membership of the Traveller/Settled Community

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Traveller	0.7	1.1	441.6	19.76	432.3	24.36	436.1	21.60
Settled	93.2	98.9	505.2	2.42	517.7	2.56	507.7	2.71
Missing	6.1	0.0	473.5	9.54	491.6	11.83	478.4	11.87
All Available	93.9	100.0	496.9	3.59	527.2	3.63	502.9	4.37

Mean Score Differences (Reference Category: Settled)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Traveller–Settled	-63.6	19.51	**-108.2**	**-19.1**	-85.4	24.22	**-140.7**	**-30.1**	-71.6	21.49	**-120.7**	**-22.5**
Missing–Settled	-31.7	9.50	**-53.4**	**-10.0**	-26.0	11.80	-53.0	0.9	-29.2	12.08	**-56.8**	**-1.6**

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Home Background

In this section, several aspects of student home background are examined. These include parental occupation and education; economic, social and cultural status; family structure and size; home educational resources; and the number of books in the home.

Parental Occupation/Socioeconomic Status (SES)

Students were asked to indicate their mother's and father's main occupations, and to give a brief description of the nature of their work. These responses were categorised according to the *International Standard Classification of Occupation (ISCO)* system. The resulting ISCO categories were transformed into an *International Socioeconomic Index* (ISEI) according to a methodology developed by Ganzeboom, de Graaf and Treiman (1992) and Ganzeboom and Treiman (1996). This 75-point occupational scale is continuous. For the analysis presented in this report, the ISEI score corresponding to the highest occupation of either parent (where available) was taken as an indicator of student socioeconomic status (SES).[24] Students were then categorised as high, medium and low, depending on the third of the distribution of SES scores into which the highest of their parents' occupations fell (Table 4.8).

Students with high-SES scores significantly outperformed students with medium and low scores in the three PISA domains. High-SES students had mean scores in mathematics, reading and science that were higher than those of low-SES students by about seven-tenths of a standard deviation in each domain.

[24] One category of occupation not accounted for by ISEI is that of homemaker. Homemaker was therefore categorised as 'missing' by the PISA international consortium. In cases where one parent's occupation was categorised as missing, and the other represented a value on the ISEI scale, the available scale value was taken as the student's SES score.

The OECD has compared the score differences on the mathematics scale for students in each country as one moves up one standard deviation on the ISEI index (OECD, 2004b, Table 4.3a). On average across the OECD, the score increase is 33.7 points. In Ireland, the score increase is slightly smaller (27.4 points), indicating slightly less disparity in the achievements of high and low SES students compared to the OECD average.

Table 4.8. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Parental Occupation/Socioeconomic Status (SES)

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Low	31.0	32.4	473.5	3.46	484.3	3.85	470.8	4.11
Medium	33.6	35.2	506.1	2.50	521.6	2.88	509.6	2.68
High	31.1	32.5	535.7	3.42	547.8	3.38	542.5	3.49
Missing	4.3	0.0	452.0	14.97	459.2	18.3	454.6	17.82
All Available	95.7	100.0	505.1	2.33	518.0	2.53	507.7	2.60

Mean Score Differences (Reference Category: High)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Low–High	-62.2	4.70	**-73.7**	**-50.7**	-63.5	5.09	**-75.9**	**-51.1**	-71.7	5.23	**-84.4**	**-58.9**
Medium–High	-29.57	4.09	**-39.6**	**-19.58**	-26.2	3.98	**-35.9**	**-16.5**	-32.9	4.18	**-43.1**	**-22.7**
Missing–High	-83.6	15.15	**-120.7**	**-46.6**	-88.5	18.78	**-134.5**	**-42.6**	-87.9	18.26	**-132.6**	**-43.3**

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \le .05$) are highlighted in bold.

Parental Educational Attainment

Students were asked to indicate the level of schooling completed by each of their parents. The higher of the two (where applicable) was then taken as a measure of parental educational attainment. Response categories were: (1) no formal or primary-level education (2) lower second level (Junior Cycle); (3) upper second level (Senior Cycle) or some further (non-tertiary) education; (4) tertiary certificate/diploma; and (5) tertiary degree. This classification is compatible with the *International Standard Classification of Education* (ISCED) (OECD, 1999) (see Appendix B, Section B.2).

Less than 6% of students had parents whose highest level of education was primary, while a further 9.7% had parents whose highest level was lower secondary. For 43.3% of students, the highest level of education attained by either parent was upper secondary/non-tertiary. Almost equal proportions (around 20%) had parents whose highest level of education was either a third level certificate/diploma, or a third level degree. Table 4.9 shows that higher levels of parental education are associated with higher achievement in all three domains. The difference in mean achievement between students of parents with upper secondary/non-tertiary education and primary education exceeds one half of a standard deviation in each domain. Similarly, the mean score difference between students whose parents have a tertiary degree and those whose parents' highest level of education is upper secondary exceeds one-third of a standard deviation in each domain.

Table 4.9. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Parental Educational Attainment

	Frequencies (All Domains)		Maths (Comb.)		Reading		Science	
	%T	%A	Mean	SE	Mean	SE	Mean	SE
None/primary	5.6	5.7	443.1	5.94	460.3	7.08	445.4	7.42
Lower second level	9.7	10.0	476.7	4.65	492.4	5.09	472.6	6.06
Upper sec level/n tert	43.3	44.4	498.0	2.53	512.3	2.80	500.7	2.86
Tertiary cert/diploma	19.6	20.1	514.5	3.94	523.8	3.94	519.8	4.30
Tertiary degree	19.3	19.8	536.5	4.20	546.2	4.14	538.5	4.44
Missing	2.5	0.0	470.8	21.98	481.0	28.1	478.9	25.74
All Available	97.5	100.0	503.7	2.40	516.4	2.57	506.1	2.66

Mean Score Differences (Reference Category: Upper secondary/non-tertiary)

	Maths (Comb.)				Reading				Science			
	Diff	SED	CI95L	CI95U	Diff	SED	CI95L	CI95U	Diff	SED	CI95L	CI95U
None/prim–u.sec/n-tert	-54.9	6.01	**-70.7**	**-39.0**	-52.1	6.89	**-70.2**	**-33.9**	-55.4	7.14	**-74.2**	**-36.5**
Lower sec–u.sec/n-tert	-21.3	5.05	**-34.6**	**-7.9**	-19.9	5.53	**-34.5**	**-5.3**	-28.1	6.49	**-45.2**	**-11.0**
Tert cert/dip-u.sec/n-tert	16.5	4.14	**5.6**	**27.4**	11.5	3.87	**1.2**	**21.7**	19.1	4.50	**7.2**	**31.0**
Tert deg–u.sec/n-tert	38.5	4.52	**26.6**	**50.5**	33.9	4.35	**22.4**	**45.4**	37.8	5.08	**24.4**	**51.2**
Missing–u.sec/n-tert	-27.1	22.12	-85.5	31.2	-31.4	28.44	-106.4	43.7	-21.8	25.79	-89.9	46.2

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Economic, Social and Cultural Status (ESCS)

An alternative measure of socioeconomic status, the economic, social and cultural status index, was computed for Ireland and for other countries in PISA 2003 by the OECD, using three variables related to SES: highest parental educational attainment (converted to number of years of education according to the ISCED classification), highest level of parental occupation (ISEI score), and the number of home possessions, including cultural possessions and books in the home. The OECD mean score on the resulting weighted likelihood estimate (WLE) scale was set at zero and the standard deviation at 1. The Irish mean on this measure, at –0.08, is a little below the OECD mean. For analytic purposes, approximately equal proportions of students for whom data were available were categorised as having high, medium or low ESCS (Table 4.10). Students with high ESCS scores outperformed students with low or medium scores in all three domains, with differences close to one standard deviation between the high and low groups and around two-fifths of a standard deviation between the medium and high groups. In the OECD (2004b) report on PISA 2003, the average percent of variance in mathematics achievement explained by student ESCS is 22.1. In Ireland, it is somewhat lower, at 16.3%, indicating that ESCS is less strongly predictive of achievement outcomes in Ireland compared to the OECD average.

Table 4.10. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Economic, Social and Cultural Status

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Low	32.9	33.4	463.0	3.20	474.8	3.67	460.3	3.77
Medium	32.8	33.3	506.5	2.64	519.7	2.79	509.0	2.97
High	32.7	33.2	540.0	3.17	552.6	3.37	547.3	3.28
Missing	1.5	0.0	486.5	30.1	506.1	36.4	501.4	33.92
All Available	98.5	100.0	503.1	2.42	515.6	2.62	505.5	2.68

Mean Score Differences (Reference Category: High)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Low–High	-76.9	4.52	**-88.0**	**-65.9**	-77.8	5.10	**-90.3**	**-65.4**	-86.9	4.99	**-99.1**	**-74.7**
Medium–High	-33.4	3.92	**-43.0**	**-23.9**	-32.9	4.08	**-42.9**	**-22.9**	-38.2	4.47	**-49.1**	**-27.3**
Missing–High	-53.5	30.16	-127.2	20.3	-46.5	36.73	-136.3	43.3	-45.9	34.22	-129.5	37.8

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Family Structure: Household Composition

Students were asked to provide information about household composition (i.e., who usually lives at home with them). Initially, students were classified as belonging to either lone-parent, nuclear (both biological parents) or mixed (one biological parent, one step-parent) families, or coming from other family categories (e.g., one grandparent, one parent). As just 3% of students were classified as 'mixed', and did not differ in achievement from those in the 'nuclear' category, mixed and nuclear were collapsed to form a single category – dual-parent families. Around 2% of students indicated that they were living in 'other' household compositions. These comprised a range of types and the achievements of these students did not differ from those with missing data for this variable. Therefore, 'other' was combined with 'missing'.[25]

As indicated in Table 4.11, 15.1% of students lived in lone-parent families, while 81.3% lived in dual-parent families. Students in lone-parent families performed significantly less well than those in dual-parent families on all three domains, with differences of around two-fifths of a standard deviation (33.6 points) in mathematics, one-third (30.4 points) in reading, and three-tenths (28.6 points) in science. Differences between students in lone-parent families, and students for whom information on this aspect of family structure was unavailable or who lived in 'other' family types (i.e., those categorised as 'missing'), are not significant.

The OECD (2004b) has computed the increased chance of students in lone-parent families scoring in the bottom quarter of the national mathematics achievement distribution. In Ireland, this value is 1.6. It is statistically significant, and is exceeded only in Belgium and the USA, suggesting that students in Ireland from lone-parent families are particularly at risk of low achievement compared with their counterparts in most other countries. Indeed, in 16 participating countries, students from lone-parent families were no more at risk than those in dual-parent families of scoring in the bottom quarter of the national mathematics achievement distribution.

[25] This categorisation differs very slightly to that used by the OECD, where 'other' family types were combined with dual-parent family types.

Table 4.11. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Lone-Parent Status

	Frequencies (All Domains)		Maths (Comb.)		Reading		Science	
	%T	%A	Mean	SE	Mean	SE	Mean	SE
Lone-parent	15.1	15.7	475.3	4.25	490.8	4.76	481.5	4.64
Dual-parent	81.3	84.3	508.9	2.52	521.1	2.62	510.1	2.78
Missing/Other	3.6	0.0	481.5	14.6	491.4	17.4	498.4	15.72
All Available	96.4	100.0	503.6	2.47	516.4	2.64	505.7	2.70

Mean Score Differences (Reference Category: Lone-Parent Status)

	Maths (Comb.)				Reading				Science			
	Diff	SED	CI95L	CI95U	Diff	SED	CI95L	CI95U	Diff	SED	CI95L	CI95U
Dual–Lone	33.6	4.20	**24.0**	**43.2**	30.4	4.48	**20.1**	**40.6**	28.6	4.64	**18.0**	**39.2**
Miss/Oth–Lone	6.2	15.57	-29.3	41.8	0.6	18.36	-41.3	42.6	16.9	16.58	-21.0	54.7

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Family Structure: Number of Siblings

Students in Ireland reported a mean of 2.7 siblings (SD = 2.08). Table 4.12 shows the numbers of siblings in students' families, and compares the students' mean scores according to the number of siblings they have. Almost 11% of students reported that they did not have any brothers or sisters; 16.9% reported having one sibling, and the modal value, two siblings, was reported by 25.6% of students. Close to half of the students (44.7%) reported having three or more siblings. Students with no siblings performed significantly less well than those with one sibling in two of the three domains (mathematics and science). Students with three siblings, and those with four or more, also performed significantly less well in all three domains in comparison with students with one sibling. For example, on the combined mathematics scale, students with four or more siblings had a score that was over two-fifths of a standard deviation (38.8 points) below that of students with one sibling.

Home Educational Resources

Students were asked to indicate whether they had each of the following educational resources in their homes: a desk for study, a quiet place to study, and books to help with school work.[26] Students with none or one of these resources (16.3%) were categorised as having low home educational resources, those with any two (25.3%) as medium, and those with all three (57.3%) as high (Table 4.13). Across all three assessment domains, the mean score of students in the high category is significantly higher than that of students in the medium and low categories, with differences between students with high and low levels of resources exceeding one-half of a standard deviation in each domain.

[26] The OECD version of the home educational resources variable (as used in the construction of the economic, social and cultural status variable reported in Table 4.10 in this chapter) also includes a calculator and a dictionary; however, these were not included in the analysis here due to ceiling effects (93.4% of Irish students had both a dictionary and a calculator available to them at home).

Table 4.12. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Number of Siblings

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
0 Sibling	10.6	10.8	501.2	4.67	519.6	5.41	507.3	4.99
1 Sibling	16.9	17.3	521.8	3.65	535.5	4.21	525.9	4.08
2 Siblings	25.6	26.1	512.3	3.33	525.1	3.38	515.6	3.67
3 Siblings	20.7	21.2	499.1	3.63	510.9	3.80	500.4	3.81
≥4 Siblings	24.0	24.5	483.0	3.89	492.3	4.00	482.1	4.27
Missing	2.2	0.0	505.0	19.0	525.1	23.7	519.7	21.52
All Available	97.8	100.0	502.8	2.44	515.3	2.62	505.1	2.67

Mean Score Differences (Reference Category: One)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
None–One	-20.62	5.35	**-34.733**	**-6.5**	-15.83	6.34	-32.56	0.9	-18.64	5.79	**-33.92**	**-3.4**
Two–One	-9.5	3.75	-19.4	0.4	-10.3	4.54	-22.3	1.6	-10.3	4.53	-22.2	1.6
Three–One	-22.7	4.55	**-34.7**	**-10.7**	-24.5	4.67	**-36.84**	**-12.2**	-25.52	4.75	**-38.1**	**-13.0**
≥Four–One	-38.8	4.94	**-51.8**	**-25.8**	-43.2	5.48	**-57.65**	**-28.7**	-43.83	5.57	**-58.5**	**-29.1**
Missing–One	-16.8	19.03	-67.0	33.4	-10.4	24.05	-73.8	53.1	-6.2	21.91	-64.1	51.6

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Table 4.13. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Educational Resources in the Home

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Low	16.3	16.5	472.1	4.15	478.7	5.37	466.3	4.77
Medium	25.3	25.6	492.0	2.98	504.9	2.98	491.0	3.18
High	57.3	57.9	516.2	2.84	530.1	3.00	522.3	3.13
Missing	1.2	0.0	511.9	26.54	539.2	32.4	536.2	27.91
All Available	98.8	100.0	502.7	2.42	515.2	2.61	505.0	2.66

Mean Score Differences (Reference Category: High)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Low–High	-44.0	4.57	**-55.2**	**-32.8**	-51.5	5.90	**-65.9**	**-37.1**	-56.0	5.13	**-68.5**	**-43.4**
Medium–High	-24.1	3.37	**-32.4**	**-15.9**	-25.2	3.37	**-33.5**	**-17.0**	-31.2	3.94	**-40.9**	**-21.6**
Missing–High	-4.3	26.67	-69.5	60.9	9.1	32.85	-71.2	89.4	14.0	28.10	-54.7	82.7

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Books in the Home

Students were asked to indicate the number of books at home using the following scale: 0-10, 11-25, 26-100, 101-200, 201-500, and more than 500. Approximately one-third of students reported that, excluding magazines, schoolbooks and newspapers, they had between 26 and 100 books; 14.8% said they had between 11 and 25; while 10.3% reported having fewer than ten (Table 4.14). Students with between 26 and 100 books achieved mean scores that were significantly higher by over one-half of a standard deviation in each domain (ranging from 50.9 in mathematics to 69.3 points in science) than those of students with 0-10 books. Students with between 26 and 100 books did less well in all three domains than students with a higher number of books. Indeed, the mean difference in achievement between those with more than 500 books and those with 26-100 is three-fifths of a standard deviation (56.0 points) for science and about one-half of a standard deviation for mathematics and reading literacy (45.5 and 47.9 points, respectively).

Table 4.14. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Number of Books in the Home

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
0 to 10 Books	10.3	10.5	446.8	5.53	454.6	6.80	434.1	5.57
11 to 25 Books	14.8	15.1	472.2	4.12	478.9	3.87	470.0	4.54
26 to100 Books	32.8	33.5	497.6	2.82	513.6	2.93	503.4	3.10
101 to 200 Books	19.5	19.9	529.6	3.25	544.1	3.06	532.8	3.32
201 to 500 Books	13.4	13.6	533.7	4.44	543.0	4.47	537.0	4.82
500 and more Books	7.3	7.4	543.2	5.09	561.5	5.43	559.4	6.01
Missing	2.1	0.0	489.5	20.66	500.1	26.8	492.1	25.87
All Available	97.9	100.0	503.1	2.43	515.8	2.62	505.7	2.68

Mean Score Differences (Reference Category: 26 to 100 Books)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
0 to 10–26 to 100	-50.9	5.68	**-66.2**	**-35.5**	-59.0	6.73	**-77.2**	**-40.7**	-69.3	5.54	**-84.3**	**-54.3**
11 to 25–26 to 100	-25.5	3.97	**-36.2**	**-14.7**	-34.7	3.99	**-45.5**	**-23.9**	-33.4	4.73	**-46.2**	**-20.6**
101 to 200–26 to 100	32.0	4.32	**20.3**	**43.7**	30.5	3.60	**20.8**	**40.3**	29.4	3.94	**18.7**	**40.0**
201 to 500–26 to 100	36.0	4.55	**23.7**	**48.3**	29.4	5.06	**15.7**	**43.1**	33.6	5.28	**19.3**	**47.9**
>500–26 to 100	45.5	5.55	**30.5**	**60.5**	47.9	5.88	**32.0**	**63.8**	56.0	6.46	**38.5**	**73.5**
Missing–26 to 100	-8.2	20.73	-64.3	48.0	-13.5	27.20	-87.1	60.2	-11.3	26.25	-82.4	59.7

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Student Out-of-School Activities

In this section, performance in mathematics, reading, and science is related to student out-of-school activities in the areas of homework practices and engagement in leisure reading of fiction and e-mail/webpages.

Homework and Study – Total Time

Students were asked to provide information about the total amount of time they typically spend on homework during the week, including weekends. Students reported spending an average of 7.7 hours (SD = 5.69) on homework/study per week across all subjects. Responses were categorised into low, medium and high.[27] Students who spent the least amount of time on homework (those in the low category) achieved a significantly lower mean score in each domain than students who spend medium or high amounts of time on homework (Table 4.15). In the case of mathematics, the difference in mean achievement scores between those categorised as low and high is over two-fifths of a standard deviation (38.8 points). Close to 9% of students are missing data on this variable; their mean score on reading is significantly lower than that of pupils in the low group.

Table 4.15. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Total Time Spent on Homework/Study

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Low	30.4	34.5	484.0	4.16	502.6	4.51	484.3	4.85
Medium	35.6	38.7	516.1	3.16	533.9	3.16	519.9	3.32
High	25.4	26.9	522.7	3.58	545.1	3.29	530.7	3.87
Missing	8.5	0.0	463.7	7.43	473.0	9.25	466.2	9.33
All Available	91.5	100.0	508.2	2.36	521.9	2.57	509.8	2.62

Mean Score Differences (Reference Category: Low)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Medium–Low	32.2	4.61	**20.9**	**43.4**	31.2	5.07	**18.8**	**43.6**	35.6	5.29	**22.7**	**48.5**
High–Low	38.8	5.08	**26.4**	**51.2**	42.5	4.99	**30.3**	**54.7**	46.4	5.56	**32.8**	**60.0**
Missing–Low	-20.2	8.38	-40.7	0.3	-29.6	9.76	**-53.5**	**-5.8**	-18.1	10.23	-43.1	6.9

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Homework and Study – Mathematics

Students reported spending an average of 2.8 hours (SD = 2.38) on mathematics homework/study per week. Again, students were categorised as high, medium, or low, depending on how much time they reported spending on maths homework/study.[28] In this instance, the achievement gap between students in the high and low categories was smaller than the gap between students in the medium and low categories, arising from the somewhat lower mean score of students in the high homework category relative to their counterparts in the medium category (Table 4.16). Just over 9% of students are missing data on this variable, and their mean score on reading is significantly lower than the mean score of students in the low group.

[27] The mean time spent on homework/study (based on student reports) is 2.2 hours (SD = 1.19) for the low group, 7.4 hours (SD = 1.94) for the medium group, and 15.3 hours (SD = 4.08) for the high group.

[28] The mean weekly time spent on mathematics homework/study (based on student reports) was 0.8 hours (SD = 0.40) for the low group, 2.5 hours (SD = 0.50) for the medium group, and 5.8 hours (SD = 2.47) for the high group.

Table 4.16. Mean Combined Mathematics Scores of Irish Students, and Mean Score Differences, by Time Spent on Mathematics Homework

	%T	*%A*	*Mean*	*SE*
Low	29.4	32.5	491.0	3.55
Medium	36.1	39.9	519.7	3.01
High	25.0	27.7	512.0	3.05
Missing	9.4	0.0	450.8	7.86
All Available	90.6	100.0	508.3	2.37

Mean Score Differences (Reference Category: Low)

	Maths (Comb.)			
	Diff	*SED*	*CI95L*	*CI95U*
Medium–Low	28.7	4.08	**18.7**	**38.7**
High–Low	21.0	4.10	**11.0**	**31.1**
Missing–Low	-40.2	8.06	**-59.9**	**-20.5**

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Leisure Reading – Frequency of Reading Fiction

Students were asked to respond to questions about the frequency with which they engaged in leisure reading of fiction out of school time. Responses were recorded on a 5-point scale ranging from ‘a few times a week’ to ‘hardly ever or never’. Almost one-quarter of students reported that they hardly ever or never read fiction, while the same proportion said that they read fiction a few times a year (Table 4.17). In contrast, 12.9% reported reading fiction several times a week. Those who read fiction a few times a year had a mean reading score that was over two-fifths of a standard deviation (38.1 points) higher than those who hardly ever or never read fiction. The difference in reading achievement between those who read fiction several times a week and those who hardly ever or never did so was over four-fifths of a standard deviation (76.5 points).

Similar patterns of achievement were also observed in mathematics and science for students engaging in varying amounts of reading fiction.

Leisure Reading – Frequency of Reading E-mails and Webpages

In a related question, students were asked about the frequency with which they read e-mails and webpages. About three in ten students reported that they never read e-mails or webpages (Table 4.18). The mean reading achievement score of these students was over one-quarter of a standard deviation (23.4 points) lower than the mean of students who read these materials a few times a year. The mean reading scores of students reading emails/webpages once a month, once a week, or several times a week are similar to one another, in contrast to the more linear pattern of increase associated with the frequency of reading fiction. Similar patterns were observed for mathematics and science. Students missing data on this variable (6.3%) had significantly lower mean scores in reading than those in the reference group. Students who hardly ever or never read e-mails and webpages also had lower mean scores in mathematics and science, compared with those who read these items more often.

Table 4.17. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Frequency of Reading Fiction

	Frequencies (All Domains)		Maths (Comb.)		Reading		Science	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Hardly ever/Never	24.5	26.0	476.3	3.14	482.4	3.31	471.0	3.55
Few times a year	24.5	25.9	510.4	2.82	520.5	3.23	509.4	3.40
Once month	21.2	22.5	509.0	3.64	528.3	3.61	514.5	4.00
Once week	11.3	11.9	517.7	4.63	537.4	5.07	525.3	5.17
Sev times week	12.9	13.7	540.7	5.63	558.9	5.45	554.0	5.62
Missing	5.6	0.0	445.3	12.66	445.7	15.89	452.2	14.13
All Available	94.4	100.0	506.2	2.33	519.6	2.48	508.5	2.61

Mean Score Differences (Reference Category: Hardly ever/Never)

	Maths				Reading				Science			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Few t yr–Hardly	34.2	3.58	**24.7**	**43.6**	38.1	4.41	**26.5**	**49.8**	38.4	4.73	**25.9**	**50.9**
Once mth–Hardly	32.7	4.23	**21.5**	**43.8**	45.9	4.77	**33.3**	**58.5**	43.5	4.92	**30.6**	**56.5**
Once wk–Hardly	41.4	4.83	**28.7**	**54.2**	55.0	5.36	**40.9**	**69.2**	54.3	5.19	**40.6**	**68.0**
Sev t wk–Hardly	64.4	6.06	**48.4**	**80.3**	76.5	6.00	**60.7**	**92.3**	83.0	6.44	**66.0**	**100.0**
Missing–Hardly	-31.0	12.77	-64.7	2.7	-36.7	15.90	-78.6	5.3	-18.8	14.46	-56.9	19.4

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Table 4.18. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Frequency of Reading E-mails/Webpages

	Frequencies (All Domains)		Maths (Comb.)		Reading		Science	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Hardly ever/Never	29.3	31.3	480.6	2.88	494.1	3.03	478.2	3.43
Few times a year	10.0	10.6	506.2	4.94	517.6	5.23	502.6	6.13
Once month	17.1	18.3	513.2	3.57	529.2	3.83	518.1	4.17
Once week	18.9	20.1	522.3	3.50	535.2	3.46	527.7	3.74
Sev times week	18.4	19.7	524.9	4.06	537.3	4.25	532.2	4.37
Missing	6.3	0.0	450.1	10.98	451.5	13.54	456.8	12.21
All Available	93.7	100.0	506.4	2.31	519.8	2.47	508.7	2.59

Mean Score Differences (Reference Category: Hardly ever/Never)

	Maths				Reading				Science			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Few t yr–Hardly	25.6	4.84	**12.8**	**38.4**	23.4	5.27	**9.5**	**37.3**	24.4	5.94	**8.7**	**40.1**
Once mth–Hardly	32.7	4.01	**22.1**	**43.3**	35.0	4.37	**23.5**	**46.6**	39.9	5.22	**26.1**	**53.7**
Once wk–Hardly	41.8	3.88	**31.6**	**52.0**	41.0	3.94	**30.6**	**51.4**	49.5	4.58	**37.4**	**61.6**
Sev t wk–Hardly	44.3	4.60	**32.2**	**56.4**	43.2	4.98	**30.0**	**56.3**	54.1	5.10	**40.6**	**67.5**
Missing–Hardly	-30.44	10.894	**-59.2**	**-1.7**	-42.7	13.60	**-78.6**	**-6.8**	-21.43	12.25	-53.76	10.89

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Academic Characteristics and Behaviour

This section explores a number of student academic characteristics and behaviours as they relate to performance in mathematics, reading and science. These included students' current grade level, the level at which they had taken/expected to take the Junior Certificate Examination in mathematics, English, and science, and the frequency with which they were absent from school.

Current Grade (Year) Level

Since PISA involved an age-based sample, a comparison was made between the mean scores of students in the Irish sample who were in the Second, Third, Fourth (Transition), and Fifth years at the time of the assessment. Sixty-one percent of students were in Third Year (the modal grade level). Just 2.8% were in Second year, while 16.7% and 19.6% were in the Fourth and Fifth years respectively (Table 4.19). Students in Third year outperformed their counterparts in Second year in all three domains, with the differences close to or exceeding one standard deviation (85.5, 96.3, and 93.5 points, in mathematics, reading, and science, respectively). Third year students performed less well than students in Fourth and Fifth years. Whereas the mean score differences in favour of Fourth years over Third years ranged from almost three-fifths of a standard deviation in mathematics (50.6 points) and science (54.5 points) to two-thirds (59.2 points) in reading, those in favour of Fifth years over Third years were smaller – about one quarter of a standard deviation in mathematics (22.8 points) and science (24.7 points), and about one-third (28.0 points) in reading. Hence, as in PISA 2000, Fifth year students did less well than their counterparts in Fourth year, suggesting that variables other than syllabus coverage may be implicated. Issues such as grade repetition, availability or uptake of the Transition year, and ways in which Transition year learning experiences may prepare students for the real world may also be relevant.

Table 4.19. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Current Grade (Year) Level

	Frequencies (All Domains)	*Maths (Comb.)*		*Reading*		*Science*	
	%T	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
2nd Year	2.8	406.8	9.48	406.2	10.01	400.5	9.95
3rd Year	60.9	492.3	2.97	502.8	3.23	494.1	3.30
4th Year	16.7	542.9	4.56	562.0	4.48	548.6	4.71
5th Year	19.6	515.1	5.32	530.8	4.36	518.8	5.23
All Available	100.0	502.8	2.45	515.5	2.63	505.4	2.69

Mean Score Differences (Reference Category: 3rd Year)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
2nd–3rd	-85.5	9.88	**-109.6**	**-61.3**	-96.6	10.68	**-122.7**	**-70.5**	-93.5	10.58	**-119.4**	**-67.7**
4th–3rd	50.6	5.08	**38.2**	**63.1**	59.2	4.78	**47.5**	**70.9**	54.5	5.14	**41.9**	**67.1**
5th–3rd	22.8	5.86	**8.5**	**37.1**	28.0	4.90	**16.0**	**39.9**	24.7	5.93	**10.2**	**39.2**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Absence from School

Students were asked to indicate the number of days on which they had been absent from school in the two weeks prior to the assessment. A majority (58.2%) reported attending every day, almost one-third had been absent for one or two days, while 9.7% indicated that they had been absent for three or more days (Table 4.20). In each domain, students with full attendance significantly outperformed students with lower levels of attendance.

In mathematics, the mean score of students with full attendance was about three-fifths of a standard deviation (49.7 points) higher than the mean score of students who reported three or more absences.[29]

Table 4.20. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Frequency of Absence from School (Past Two Weeks)

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
None	58.2	60.2	514.7	2.62	527.0	2.73	517.9	2.89
1 or 2	28.8	29.8	495.1	3.22	509.6	3.30	496.8	3.60
≥3	9.7	10.0	465.0	6.52	473.7	6.48	463.0	7.04
Missing	3.3	0.0	471.8	16.14	486.1	20.23	483.3	18.91
All Available	96.7	100.0	503.9	2.41	516.5	2.57	506.1	2.67

Mean Score Differences (Reference Category: None)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
1 or 2–None	-19.57	3.18	**-27.3**	**-11.8**	-17.33	3.18	**-25.1**	**-9.6**	-21.1	3.60	**-29.9**	**-12.3**
≥3–None	-49.7	6.38	**-65.3**	**-34.1**	-53.3	6.28	**-68.6**	**-37.9**	-54.9	6.80	**-71.5**	**-38.3**
Missing–None	-42.9	15.93	**-81.8**	**-3.9**	-40.9	20.00	-89.8	8.0	-34.6	18.73	-80.4	11.2

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Early School-Leaving Intent

Students who indicated that they did not intend to complete a programme leading to the Leaving Certificate Examination, together with agreement with at least one of eight statements about possible reasons for leaving school early (e.g., 'I don't like school', 'A lot of my friends are leaving school', 'I want to do an apprenticeship'), were identified as being at risk of early school leaving. [30] Based on these criteria, 20.2% of students were identified as being at risk (Table 4.21), compared with 14.0% in 2000. Mean scores favouring students intending to remain in school were greater by almost nine-tenths of a standard deviation (75.1 points) in mathematics, and almost one standard deviation in both reading and science (85.9 and 86.1 points, respectively). These differences are similar in magnitude to those observed for this variable in PISA 2000.

[29] It should be noted that the absence variable did not capture the reasons why students were absent (e.g., disengagement, illness, etc.); nonetheless, there is a high correspondence between low incidence of absence and low incidence of skipping off classes: 88.0% of students who reported missing no days also reported no incidences of skipping off classes, so the variable does appear to capture disengagement from school.

[30] In the Irish national report on PISA 2000, this variable was called 'Dropout Risk'.

Table 4.21. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Early School-Leaving Intent

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
No	78.7	79.5	518.0	2.31	532.6	2.37	522.6	2.55
Yes	20.2	20.5	442.8	3.91	446.8	4.18	436.4	3.92
Missing	1.1	0.0	526.9	22.16	554.1	27.33	548.5	25.15
All Available	98.9	100.0	502.6	2.44	515.1	2.61	504.9	2.67

Mean Score Differences (Reference Category: No)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Yes–No	-75.1	3.52	**-83.2**	**-67.1**	-85.9	3.69	**-94.3**	**-77.4**	-86.1	3.56	**-94.3**	**-78.0**
Missing–No	8.9	21.99	-41.3	59.1	21.5	27.30	-40.9	83.8	26.0	25.09	-31.3	83.3

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Calculator Usage in PISA 2003

Calculator usage was optional in the PISA 2003 assessment. Students were provided with, and could use, a calculator during the assessment, if they wished. Almost 80% of students indicated that they had used a calculator during the mathematics component of the assessment, while almost 9.7% reported that they had not (Table 4.22). In PISA 2000, calculator use was not as widespread, with just 24.2% of students reporting use. The mean score in mathematics in 2003 of students who used a calculator is just under one-quarter of a standard deviation (20.3 points) higher than the mean score of those who did not. However, it cannot be inferred that calculator use *per se* was responsible for this difference. Further, the mean mathematics achievement of about one-eighth of students (12.3%) who are missing data on this variable is significantly lower than the mean of students in the reference group.

Table 4.22. Mean Combined Mathematics Scores of Irish Students, and Mean Score Differences, by Use of Calculators in the PISA Assessment

	%T	*%A*	*Mean*	*SE*
Yes	78.0	89.0	513.1	2.27
No	9.7	11.0	492.8	5.42
Missing	12.3	0.0	445.9	4.88
All Available	87.7	100.0	510.8	2.26

Mean Score Differences (Reference Category: No)

	Maths (Comb.)			
	Diff	*SED*	*CI95L*	*CI95U*
Yes–No	20.3	5.33	**8.1**	**32.5**
Missing–No	-46.9	7.06	**-63.0**	**-30.8**

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Study of Science for the Junior Certificate

Students were asked to indicate whether or not they were studying, or had studied, science as a subject for the Junior Certificate Examination. Just 9.9% indicated that they did not study science (Table 4.23). These students performed significantly less well in all three PISA domains than students who studied science. In science, the mean score of students who did not study science was two-thirds of a standard deviation (61.7 points) lower than the mean score of students who reported studying the subject. The magnitude of the difference is similar for mathematics (59.7 points), while it is just under half a standard deviation (44.7 points) for reading literacy.

Table 4.23. Mean Combined Mathematics, Reading, and Science Scores of Irish Students, and Mean Score Differences, by Study of Science at Junior Certificate Level

	Frequencies (All Domains)		Maths (Comb.)		Reading		Science	
	%T	%A	Mean	SE	Mean	SE	Mean	SE
Yes	88.4	89.9	510.8	2.29	522.3	2.46	513.5	2.43
No	9.9	10.1	451.1	5.67	477.6	6.86	451.8	6.52
Missing	1.7	0.0	388.8	13.18	380.1	15.93	394.6	13.56
All Available	98.3	100.0	504.8	2.38	517.8	2.54	507.3	2.62

Mean Score Differences (Reference Category: Yes)

	Maths (Comb.)				Reading				Science			
	Diff	SED	CI95L	CI95U	Diff	SED	CI95L	CI95U	Diff	SED	CI95L	CI95U
No–Yes	-59.69	5.78	**-72.9**	**-46.5**	-44.69	6.90	**-60.5**	**-28.9**	-61.67	6.41	**-76.3**	**-47.0**
Missing–Yes	-122.0	13.39	**-152.6**	**-91.4**	-142.2	16.28	**-179.4**	**-105.0**	-118.9	13.76	**-150.3**	**-87.4**

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Syllabus Level of Mathematics, English and Science for the Junior Certificate

Students indicated which syllabus level they had taken, or intended to take, for the Junior Certificate Examination, in mathematics, English, and science. The actual levels taken by students in either 2002 or 2003 were obtained from the Junior Certificate Examinations database of the State Examinations Commission (93.9% of all PISA students), while students sitting the examination in a year other than 2002 or 2003 were assigned the levels indicated by them in the Student Questionnaire.

Table 4.24 shows the mean scores of students taking the examination at each syllabus level in the corresponding Junior Certificate Examination subject. In all three subject domains, there is a mean score difference of just over one standard deviation between students taking the corresponding subject at Higher and Ordinary levels. In the case of mathematics and English, there are also mean score differences of around a standard deviation between Ordinary and Foundation levels on PISA mathematics and PISA reading literacy. The mean score difference on PISA science between students taking Junior Certificate science at Ordinary level and those not taking science for the Junior Certificate is only around nine score points (around one-tenth of standard deviation) and is not statistically significant.

In terms of the cut-points associated with the proficiency levels for mathematics and reading (Chapter 3), in the case of mathematics, the mean score of Higher level students is at Level 4. Ordinary level students have a mean score at Level 2, and Foundation level students have a mean score at Level 1. In the case of reading literacy, students taking Higher level English have a mean score at the top of Level 3; those taking Ordinary at the middle of Level 2; while and Foundation level students score towards the bottom of Level 1. In science, Ordinary level students have a score (443.3) around the Irish 25th percentile (441.8), while the mean score (547.1) of students taking Higher level science is 24.8 points below the 75th percentile score (571.9) for Irish students.

Table 4.24. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Syllabus Level Taken/to be Taken, at Junior Certificate, in the Corresponding Subject Areas

	Frequencies		*Maths (Comb.)*		*Frequencies*		*Reading*		*Frequencies*		*Science*	
	%T	*%A*	*Mean*	*SE*	*%T*	*%A*	*Mean*	*SE*	*%T*	*%A*	*Mean*	*SE*
Higher	42.8	42.5	563.0	2.09	68.8	69.2	548.2	2.11	59.8	60.8	547.1	2.25
Ordinary	50.3	50.0	469.1	1.99	29.1	29.3	449.1	3.12	28.6	29.1	443.3	3.27
Foundation	6.9	6.8	385.4	5.16	1.4	1.5	355.9	14.05	—	—	—	—
Don't study	—	—	—	—	—	—	—	—	9.9	10.1	451.8	6.52
Missing	0.6	0.0	387.4	22.49	0.7	0.0	381.3	25.1	1.7	0.0	394.6	13.6
All Available	99.4	100.0	503.6	2.45	99.3	100.0	516.4	2.63	98.3	100.0	507.3	2.6

Mean Score Differences (Reference Category: Ordinary)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Hig–Ord	93.9	2.88	**86.8**	**100.9**	99.1	3.762	**89.9**	**108.3**	103.8	3.971	**94.1**	**113.5**
Found–Ord	-83.7	5.53	**-97.2**	**-70.2**	-93.3	14.39	**-128.4**	**-58.1**	—	—	—	—
Don't st.–Ord	—	—	—	—	—	—	—	—	8.5	7.3	-9.3	26.4
Miss–Ord	-81.7	22.58	**-136.9**	**-26.5**	-67.9	25.24	**-129.6**	**-6.2**	-48.7	13.95	**-82.8**	**-14.5**

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Students' Attitudes Towards Mathematics

Data were obtained on a range of variables designed to tap into students' interest in mathematics, their confidence in tackling mathematics problems, and their mathematics self-concept. Here, the responses of Irish students on two variables – perceived self-efficacy in mathematics, and self-reported anxiety about mathematics – are examined in terms of their associations with performance on PISA mathematics. Gender differences on the self-efficacy and anxiety variables are also considered.

Self-Efficacy in Mathematics

Self-efficacy in mathematics was measured by asking students to rate their confidence in solving a number of mathematical tasks such as calculating how much cheaper a TV set would be after a 30% discount, solving a simple equation for x (e.g., $3x + 5 = 17$), solving a more complex equation (e.g., $2(x + 3) = (x + 3)(x – 3)$), and calculating the petrol consumption rate of a car. The tasks were rated using a 4-point scale ranging from 'very confident' to 'not at all confident'. A weighted likelihood estimate composite was constructed with an OECD mean of zero, and a standard deviation of 1. Students were categorised as low, medium, or high, based on the degree of confidence they reported in their ability to solve the problems.

Students who reported high self-efficacy in mathematics achieved a significantly higher mean score in the domain than students with medium or low self-efficacy (Table 4.25). Indeed, the difference in favour of high self-efficacy students over low self-efficacy students was over one and one-quarter standard deviations (108.5 points).

In international terms, Irish students had a mean self-efficacy score of –0.03, which is not significantly different from the OECD country average of zero, and is significantly higher than the mean scores of students in some high-scoring countries in mathematics such as Korea (–0.42) and Japan (–0.53). Male students had significantly higher self-efficacy scores than females in all countries. In Ireland, the difference was close to the OECD average difference of one-third of a standard deviation.

Table 4.25. Mean Combined Mathematics Scores of Irish Students, and Mean Score Differences, by Perceived Self-Efficacy in Mathematics

	%T	*%A*	*Mean*	*SE*
Low	30.4	30.9	450.9	2.50
Medium	38.9	39.6	502.8	2.87
High	29.0	29.5	559.4	2.70
Missing	1.6	0.0	465.7	34.15
All Available	98.4	100.0	503.4	2.39

Mean Score Differences (Reference Category: High)

	Maths (Comb.)			
	Diff	*SED*	*CI95L*	*CI95U*
Medium–Low	-108.5	3.49	**-117.0**	**-100.0**
High–Low	-56.6	3.24	**-64.5**	**-48.7**
Missing–Low	-93.7	34.00	**-176.9**	**-10.5**

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Anxiety About Mathematics

Students were asked to respond to a number of statements designed to measure their general concerns, if any, about their achievement in mathematics. These statements included, ‘I get very nervous about doing mathematics problems’, and ‘I worry that I will get poor grades in mathematics’. Students indicated their agreement with the statements on a 4-point scale ranging from ‘strongly agree’ to ‘strongly disagree’. As in the case of the self-efficacy variable, a composite was constructed with an OECD mean of zero, and a standard deviation of 1 (with positive scores indicating greater anxiety), and students were categorised as low, medium, or high.

Students with high levels of anxiety achieved a mean score in mathematics that was four-fifths of a standard deviation (68.7 points) lower than the mean score of students with low anxiety (Table 4.26). The difference in favour of those with medium anxiety over those with high anxiety was two-fifths of a standard deviation (34.5 points).

Ireland's mean score on the anxiety scale (–0.07) was a little below the OECD average (0.0). Mean anxiety was considerably lower in a number of countries including the Netherlands (–0.38), Denmark (–0.46), and Sweden (–0.49) (OECD, 2004b, Table 3.8). In all countries except Poland and Serbia, male students reported significantly less anxiety about mathematics than females. The difference in Ireland (about one-sixth of a standard deviation) is smaller than the OECD average (one-quarter of a standard deviation), but statistically significant nonetheless.

Table 4.26. Mean Combined Mathematics Scores of Irish Students, and Mean Score Differences, by Anxiety About Mathematics

	%T	*%A*	*Mean*	*SE*
Low	30.7	31.3	536.8	3.40
Medium	39.8	40.5	502.6	3.34
High	27.7	28.2	468.1	2.95
Missing	1.8	0.0	459.9	32.41
All Available	98.2	100.0	503.6	2.37

Mean Score Differences (Reference Category: Low)

	Maths (Comb.)			
	Diff	*SED*	*CI95L*	*CI95U*
Medium–Low	-34.2	4.37	**-44.9**	**-23.5**
High–Low	-68.7	4.38	**-79.4**	**-58.0**
Missing–Low	-77.0	32.16	-155.6	1.6

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

SCHOOL CHARACTERISTICS AND ACHIEVEMENT

Relationships between school characteristics and students' achievement in mathematics, reading literacy and science are examined in this section. Two broad categories of school-level variables are described: school structure and composition, and school climate and resources. The data were derived from three sources – principals' responses on the School Questionnaire, students' responses on the Student Questionnaire, and the post-primary schools' database of the Department of Education and Science. Variables that were originally collected or computed at the student level (disciplinary climate; economic, socials and cultural status) were aggregated to the school level and were then disaggregated to the student level, so that each student was assigned the mean value for his/her school.

School Structure and Composition

In this section, the following aspects of school structure and composition are considered: school size (stratum); school sector; school disadvantaged status; economic, social and cultural status; gender composition; and Junior Certificate Examination fee waiver entitlement.

School Size (Stratum)

For sampling purposes, schools were categorised according to the number of 15-year-olds enrolled as follows: small (1-40 students), medium (41-80), and large (81 or more). This created an explicit stratifying variable called school size (see Chapter 1). The majority of students (67.7%) attended large schools, while less than 5% attended small schools (Table 4.27). These are weighted percentages which are representative of the student population.

Students attending large schools had significantly higher mean scores in all three assessment domains than students attending medium-sized schools, although differences tended to be small – about one-fifth of a standard deviation in each case (Table 4.27). Although the mean scores of students in small schools were lower than those of students in large schools in all three domains, differences are not statistically significant at the .05 level, except in reading, where the difference is over three-fifths of a standard deviation (54.0 points). In science, the difference in favour of those attending large schools was found to be borderline significant (Bonferroni 90%; confidence interval: –91.1 to –1.8). These findings should be interpreted with reference to the relatively small proportion of students in the small school stratum and hence the large standard errors associated with their mean scores.

Table 4.27. Mean Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by School Size (Stratum)

	Frequencies (All Domains)	*Maths (Comb.)*		*Reading*		*Science*	
	%T	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Small	4.2	471.2	23.29	469.6	21.03	466.4	22.28
Medium	28.1	491.5	5.25	502.7	5.96	493.2	5.84
Large	67.7	509.5	2.65	523.6	2.88	512.9	2.96
All Available	100.0	502.8	2.45	515.5	2.63	505.4	2.69

Mean Score Differences (Reference Category: Large)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Small–Large	-38.3	23.42	-91.8	15.2	-54.0	21.28	**-102.6**	**-5.4**	-46.4	22.42	-97.7	4.8
Med–Large	-18.0	5.90	**-31.4**	**-4.5**	-21.0	6.64	**-36.1**	**-5.8**	-19.7	6.58	**-34.7**	**-4.6**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

School Sector

Schools were categorised for the study according to whether they were secondary, vocational, or community/comprehensive, creating the implicit stratifying variable, school sector. Each student was assigned the sector corresponding to the school s/he attended. The majority of students (61.0%) attended secondary; 21.7% attended vocational schools; and 17.3% attended community/comprehensive schools (Table 4.28).

Table 4.28. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by School Sector

	Frequencies (All Domains)	Maths (Comb.)		Reading		Science	
	%T	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Com./Compr.	17.3	497.6	5.05	505.9	5.28	498.7	5.69
Secondary	61.0	514.4	3.28	531.1	3.38	518.7	3.52
Vocational	21.7	474.4	5.52	479.1	6.46	473.4	6.26
All Available	100.0	502.8	2.45	515.5	2.63	505.4	2.69

Mean Score Differences (Reference Category: Secondary)

	Maths (Comb.)				Reading				Science			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Com./Compr.–Sec.	-16.9	6.01	**-30.6**	**-3.2**	-25.2	6.27	**-39.6**	**-10.9**	-20.0	6.79	**-35.5**	**-4.5**
Voc.–Sec.	-40.1	6.50	**-54.9**	**-25.3**	-52.0	7.30	**-68.7**	**-35.3**	-45.3	7.17	**-61.7**	**-28.9**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Students attending secondary schools achieved significantly higher mean scores than students attending vocational or community/comprehensive schools in the three assessment domains. Differences in achievement were smaller between secondary and community/comprehensive schools than between secondary and vocational schools. For example, the difference in mathematics between students attending secondary schools and students attending community/comprehensive schools was one-fifth of a standard deviation (16.9 points), while the difference between students in secondary schools and students in vocational schools was almost one-half of a standard deviation (40.1 points). While in PISA 2000, the difference in mathematics achievement between students in secondary and community/comprehensive schools was only statistically significant at the .10 level, the difference in 2003 was significant at the .05 level.

Disadvantaged Status

Schools were categorised for the study according to whether or not they were in the Department of Education and Science Designated Disadvantaged Area Schools Scheme. Schools in the scheme are provided with additional support, including additional teaching posts, and enhanced grants for equipment, resources, and home-school activities.

Over one-quarter of students attended schools designated as disadvantaged (Table 4.29). Students in these schools achieved mean scores in all three domains that were significantly lower than the mean scores of students attending schools that were not designated. Mean differences amount to about two-fifths of a standard deviation in each domain. In PISA 2000, mean score differences were marginally higher, at about half a standard deviation.

Table 4.29. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by School Disadvantaged Status

	Frequencies (All Domains)	*Maths (Comb.)*		*Reading*		*Science*	
	%T	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Not disadv.	74.6	512.3	2.83	524.9	2.98	515.2	2.98
Disadv.	25.4	477.0	4.79	489.8	5.67	478.6	5.50
All Available	100.0	500.3	2.71	522.6	2.77	506.4	3.43

Mean Score Differences (Reference Category: Not Disadvantaged)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Dis.–N. disadv.	-35.3	5.82	**-46.9**	**-23.7**	-35.1	6.73	**-48.5**	**-21.7**	-36.6	6.44	**-49.5**	**-23.8**

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

School Economic, Social and Cultural Status (ESCS)

As indicated earlier in this chapter, an index of the economic, social and cultural status (ESCS) of schools was derived from information obtained from the Student Questionnaire. For the purpose of the analyses presented here, scores for individual students were averaged within schools. Then each student was assigned his/her school's score, and students were categorised as attending schools with low, medium, and high ESCS, based on the 33rd and 67th percentiles of the resulting distribution.

Students attending schools with high average ESCS significantly outperformed students in schools with medium and low average ESCS values in all three assessment domains (Table 4.30). For example, the mean difference in mathematics between high and low ESCS schools is three-quarters of a standard deviation (64.3 points), and between students in high and medium schools, about three-tenths of a standard deviation (25.4 points). ESCS scores are of interest as a general, global measure of educational disadvantage, as the OECD (2004b, c) uses them to describe disadvantage at the school level in the initial report on PISA 2003.

Table 4.30. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by School Economic, Social, and Cultural Status

	Frequencies (All Domains)	*Maths (Comb.)*		*Reading*		*Science*	
	%T	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Low	32.2	468.0	4.34	476.7	5.11	467.0	4.96
Medium	34.2	506.8	3.69	520.0	4.12	509.6	3.93
High	33.5	532.3	3.91	548.1	4.38	537.9	3.97
All Available	100.0	502.8	2.45	515.5	2.63	505.4	2.69

Mean Score Differences (Reference Category: High)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Low–high	-64.3	6.00	**-78.0**	**-50.6**	-71.4	7.01	**-87.5**	**-55.4**	-70.9	6.41	**-85.5**	**-56.3**
Medium–high	-25.4	5.87	**-38.8**	**-12.0**	-28.1	6.53	**-43.0**	**-13.1**	-28.3	6.28	**-42.7**	**-14.0**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

School Gender Composition

Schools were categorised for the study into four groups based on the percentage of female 15-year-olds enrolled: no girls (0%); a small proportion of girls (0.1 to 45%); a large proportion of girls (45.1 to 99.9%) and all girls (100%)[31]. Just over one-fifth of students attended schools in which no females are enrolled, while just over one-quarter attended all girls schools (Table 4.31).

Students in all boys schools outperformed students in all girls schools by just over one-quarter of a standard deviation (22.4 points) in mathematics, and by one-eighth of a standard deviation (11.8 points) in science. In reading, students in all girls schools outperformed students in all boys schools by almost one-fifth of a standard deviation (17.7 points). Students in all boys schools performed significantly better in all three assessment domains than students in schools with 0.1 to 45% female enrolment, and students in schools with 45.1 to 99.9% female enrolment. Differences between schools with no female students and schools with 0.1 to 45% female enrolment exceeded one-half of a standard deviation in each of the three assessment domains.[32]

Table 4.31. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Percent of Female 15-Year-Old Students Enrolled

	Frequencies (All Domains)	*Maths (Comb.)*		*Reading*		*Science*	
Percent Female	*%T*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
0%	20.8	528.9	4.59	524.6	4.99	528.0	4.66
0.1-45%	23.0	480.7	6.07	485.0	7.06	479.8	6.72
45.1 - 99.9%	30.9	498.7	4.30	510.0	4.58	500.4	4.75
100%	25.3	506.6	5.59	542.4	6.05	516.2	6.31
All	100.0	502.8	2.45	515.5	2.63	505.4	2.69

Mean Score Differences (Reference Category: 0%)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
0.1-45%-0%	-48.3	7.49	**-66.6**	**-30.0**	-39.6	8.68	**-60.9**	**-18.4**	-48.2	8.38	**-68.7**	**-27.7**
45.1-99.9%-0%	-30.2	6.43	**-45.9**	**-14.5**	-14.6	6.75	**-31.1**	**1.9**	-27.7	6.80	**-44.3**	**-11.0**
100%-0%	-22.4	7.03	**-39.6**	**-5.2**	17.7	7.81	**-1.4**	**36.8**	-11.8	7.92	**-31.2**	**7.5**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals.

[31] This variable was constructed by taking the percentage of 15-year-olds enrolled in each school from the sampling frame and splitting it into four categorical groups. It is not identical to the categorical implicit stratifying variable described in Chapter 1 that has five rather than four categories since the distribution of the stratifying variable is very uneven.

[32] Although of potential policy interest, variation in gender differences across schools of differing sex composition is not examined here. This is because school sex composition is itself associated with other variables. The majority (69.0%) of participating secondary schools were single sex, while 90.0% of community/comprehensive schools were mixed sex. All vocational schools in the sample were mixed sex. Further, the percent of students in schools of varying sex composition entitled to a fee waiver for the Junior Certificate is lower in all boys (16.7%) and all girls (20.8%) schools than in schools with between 1% and 45% female students (34.5%) and those with between 46% and 99% females (29.7%).

Junior Certificate Examination Fee Waiver Entitlement

Using information from the post-primary schools database of the Department of Education and Science, an index was constructed that gave the weighted percentage of students in a school who were entitled to a fee waiver for the Junior Certificate Examination. In each school, the percentage of 15-year-old students who were entitled to the waiver was weighted by the number of students in the school who took the Junior Certificate Examination in 2002 and 2003. Each student was then assigned the value of this variable for his or her school. The school-level average for Junior Certificate Examination fee waiver is 29.8% of students (SD = 16.90%).

Students attending schools with high proportions of fee-waiver recipients did significantly less well in all three domains than students attending schools with medium or low proportions of recipients (Table 4.32). Differences of seven-tenths to four-fifths of a standard deviation (depending on the domain) between students in schools with low and high proportions of fee-waiver recipients, and of about three-tenths of a standard deviation in each domain between students attending schools with low and medium proportions of fee-waiver recipients, were observed.

Table 4.32. Mean Combined Mathematics, Reading, and Science Scores of Irish Students, and Mean Score Differences, by Percentage in School Entitled to a Junior Certificate Examination Fee Waiver

	Frequencies (All Domains)	*Maths (Comb.)*		*Reading*		*Science*	
	%T	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Low	32.8	531.2	4.42	547.7	5.06	537.3	4.50
Medium	34.2	505.9	3.44	519.3	3.46	508.3	3.62
High	33.0	471.4	4.58	479.4	5.43	470.6	5.00
All Available	100.0	502.8	2.45	515.5	2.63	505.4	2.69

Mean Score Differences (Reference Category: Low)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Medium–Low	-25.3	5.90	**-38.8**	**-11.8**	-28.4	6.41	**-43.0**	**-13.7**	-28.9	6.21	**-43.1**	**-14.8**
High–Low	-59.8	6.63	**-74.9**	**-44.6**	-68.4	7.92	**-86.5**	**-50.3**	-66.7	6.84	**-82.3**	**-51.1**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

School Climate and Resources

In this section, aspects of school climate and resources are considered: the disciplinary climate in mathematics classes, computer resources, and instructional time.

Disciplinary Climate in Mathematics Classes

Students indicated the frequency with which five events occurred during mathematics classes, including: ‘There is noise and disorder’, and ‘Students don’t listen to what the teacher says’. Frequencies were reversed so that a high frequency denotes a positive disciplinary climate, and a weighted likelihood composite estimate was created at the student level. For the purposes of the analysis reported here, disciplinary climate was averaged at the school level to obtain a measure of school disciplinary climate in mathematics classes. Each student was then assigned the disciplinary climate score for his/her school, and the resulting distribution of student scores was divided into thirds, indicating students who attended schools with high, medium and low disciplinary climate. Students attending schools with a high (positive) disciplinary climate had a mean score in mathematics that was one-half of a standard deviation (43.7 points) higher than the mean score of students attending a school with a low (negative) disciplinary climate (Table 4.33). Students in schools with a high disciplinary climate outperformed students in schools with a medium climate by almost one-quarter of a standard deviation (21.0 points). Both differences are statistically significant. Similar patterns were observed for reading and science.

Table 4.33. Mean Combined Mathematics, Reading and Science Scores of Irish Students, and Mean Score Differences, by Disciplinary Climate in Mathematics Classes

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Low	29.2	29.8	482.0	3.72	493.5	3.94	484.2	4.49
Medium	40.4	41.3	504.6	2.61	515.2	2.71	505.7	2.90
High	28.3	28.9	525.7	3.60	542.4	4.03	528.8	3.98
Missing	2.1	0.0	452.1	28.55	466.6	34.88	479.2	29.08
All Available	97.9	100.0	504.0	2.35	516.6	2.54	506.0	2.63

Mean Score Differences (Reference Category: High)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Low–High	-43.7	4.79	**-55.4**	**-32.0**	-48.87	5.04	**-61.2**	**-36.6**	-44.56	5.53	**-58.1**	**-31.1**
Medium–High	-21.0	3.30	**-29.1**	**-13.0**	-27.2	3.88	**-36.7**	**-17.7**	-23.03	3.71	**-32.1**	**-14.0**
Missing–High	-73.6	28.79	**-144.0**	**-3.2**	-75.8	35.54	-162.7	11.1	-49.6	29.41	-121.4	22.3

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Ratio of Computers to Students

The ratio of computers to 15-year-old students was derived from principals' reports on the availability of computers in their school. The number of computers available to 15-year-old students was divided by the total number of 15-year-olds in the school to yield a computer-student ratio. This was then disaggregated to the student level and the resulting distribution was divided into thirds. Data on this question were unavailable for 16.1% of students.

Students attending schools with a high computer-student ratio had mean scores in all three domains that are significantly lower than the mean scores of students attending schools with medium or low student-computer ratios (Table 4.34). Students attending schools with a low computer-student ratio outperformed students attending schools with a high ratio were by one-third of a standard deviation in mathematics and science (29.9 and 34.9 points, respectively), and by over two-fifths (38.9 points) in reading. At first glance, this finding may seem counter-intuitive, but the relationship appears to be mediated by SES. For example, the percentage of students in receipt of a Junior Certificate fee waiver in schools with low computer-student ratios (19.9%) is just over one standard deviation lower than that of schools with high computer student ratios (35.8%).[33] It is also of interest that achievement differences

[33] The school-level mean and standard deviation for Junior Certificate fee waiver are 29.8% and 16.9%, respectively.

between students in schools with low and medium computer-student ratios are small and not statistically significant.

Table 4.34. Mean Combined Mathematics, Reading, and Science Scores for Irish Students, and Mean Score Differences, by Ratio of Computers to Students

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Low	28.0	33.4	513.6	4.89	529.1	5.33	518.3	4.95
Medium	27.5	32.7	510.9	4.89	527.3	5.64	515.1	5.17
High	28.4	33.9	483.7	6.01	490.2	6.66	483.3	6.47
Missing	16.1	0.0	504.1	7.67	516.2	7.95	505.3	8.59
All Available	83.9	100.0	502.6	2.49	515.3	2.80	505.4	2.72

Mean Score Differences (Reference Category: High)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Low–High	29.9	7.61	**11.3**	**48.6**	38.9	8.37	**18.4**	**59.3**	34.9	7.98	**15.4**	**54.4**
Medium–High	27.2	8.04	**7.6**	**46.9**	37.1	8.78	**15.6**	**58.5**	31.8	8.55	**10.9**	**52.7**
Missing–High	20.4	9.94	-3.9	44.7	26.0	10.73	-0.2	52.2	22.0	11.29	-5.6	49.5

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Instructional Time

Students' responses to questions about the number of class periods in the last full week and the average length of a class period were used to calculate the average weekly total hours of total instructional time in each school. Students were assigned the average instructional time for their school, and were divided into three groups – those attending schools with high, medium, and low instructional time.[34] Data on this question are unavailable for 19.4% of students.

Those attending schools with medium instructional time had significantly higher mean achievement scores in all three assessment domains than students attending schools with low instructional time (Table 4.35). Differences were in the order of one-quarter of a standard deviation in each domain. No statistically significant differences in achievement at the .05 level were found in a comparison of the performance of students attending schools with high and low instructional time. A borderline significant difference was observed for reading, in favour of students in schools with high instructional time (Bonferroni 90% confidence interval: 1.57 to 28.86).

A measure of instructional time in mathematics only was also available. However, no statistically significant differences were found between the mean mathematics scores of students attending schools with high, medium and low amounts of instructional time.[35]

In the Irish second-level education system, one would expect only limited variation in instructional time, given the standardised school year. The variation reported here may reflect variation in the number of classes offered each week, as well as the length of class periods.

[34] The mean number of hours' instructional time per week for students in low schools is 23.3 (SD = 5.94); in medium schools, it is 28.3 (SD = 0.64); and in high schools, it is 30.7 (SD = 2.41).
[35] The mean weekly number of hours' mathematics instructional time for students in low schools is 2.4 (SD = 0.63); in medium schools, it is 3.3 (SD = 0.08); and in high schools, it is 4.2 (SD = 2.41).

Table 4.35. Mean Combined Mathematics, Reading, and Science Scores of Irish Students, and Mean Score Differences, by Minutes of Instructional Time Per Week

	Frequencies (All Domains)		*Maths (Comb.)*		*Reading*		*Science*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Low	27.6	34.2	503.4	4.71	513.4	4.65	504.2	5.12
Medium	26.5	32.8	525.7	4.81	539.7	4.82	528.9	5.10
High	26.6	33.0	512.9	3.23	528.6	4.09	514.1	3.94
Missing	19.4	0.0	457.1	5.40	467.3	6.50	463.0	6.00
All Available	80.6	100.0	513.8	2.40	527.1	2.51	515.6	2.65

Mean Score Differences (Reference Category: Low)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Medium–Low	22.3	7.21	**4.7**	**39.92**	26.3	6.96	**9.3**	**43.3**	24.7	7.65	**6.0**	**43.4**
High–Low	9.5	5.68	-4.4	23.4	15.2	6.30	-0.2	30.6	9.9	6.48	-5.9	25.7
Missing–Low	-46.3	6.91	**-63.2**	**-29.4**	-46.1	7.72	**-65.0**	**-27.3**	-41.2	7.76	**-60.2**	**-22.2**

Note. N = 3880. %T = percentage all; %A = percentage available (apply to all 3 domains); Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

CORRELATIONS BETWEEN EXPLANATORY VARIABLES AND ACHIEVEMENT

In analyses up to this point, student and school variables were split into ordered categories, and associations between each category and achievement were considered. In this section the linear associations (correlations) between the variables in their original form and the combined mathematics, reading and science scales are presented. In a few cases, linear associations between variables with a number of ordered categories (e.g., number of siblings) and achievement are presented. Inset 4.4 provides information on the computation and interpretation of the correlation coefficients reported here.

Student-Level Variables

As was the case in PISA 2000, the linear associations between variables at the student level and achievement outcomes in the three domains in 2003 are weak to moderate. The economic, social and cultural status (ESCS) of a student's family is positively associated with achievement, with correlations in the moderate or moderate/strong range (.40 for combined mathematics, .41 for reading, and .42 for science) (Table 4.36). These are stronger than the correlations between socioeconomic status and student achievement (which vary between .32 and .34), or between parent education and student achievement (.24 to .27). However, since the combined economic, social and cultural status variable consists of a number of variables which may interact with one another in complex ways, in Chapter 5, socioeconomic status and parental educational attainment are preferred for inclusion in the multilevel models.

Inset 4.4. Computation and Interpretation of Correlation Coefficients

Regardless of the strength of the relationship between two variables, it cannot be inferred that there is a causal relationship between them. One or more additional variables may contribute to the relationship.

Correlation coefficients were computed by first running linear regressions in WesVar 4.2 involving the explanatory variable and the five plausible values for the corresponding response variable (achievement in one of the assessment domains), and then obtaining the square roots of the resulting r^2 s. Since the estimated *r* coefficients might not be Normally distributed, each was transformed to a *z* score using Fisher's transformation (Schafer, 1997), and the resulting *z* scores were averaged and back-transformed. The significance of *r* was determined by computing the *t* statistic (i.e., by dividing the mean coefficient resulting from five linear regressions by its standard error); this also provides a test of linear association in the population. The corresponding *p* value was obtained from a table of critical values of *t*, using 80 degrees of freedom (an approximation based on the number of variance strata in the BRR variance estimation method).

Where the linear association between an explanatory and response variable is reported as a correlation coefficient, the following interpretation applies: a one-standard deviation increase in the explanatory variable is associated with an increase in the response (achievement) variable that is the product of its standard deviation and the correlation coefficient. Moreover, this relationship is symmetrical, implying that a standard deviation increase in the response variable is associated with an increase in the explanatory variable that is the product of its standard deviation and the correlation coefficient.

Correlation coefficients range on a scale from –1.0 to +1.0. A positive correlation indicates that an increase in the value of explanatory variable is associated with an increase in the value of the response variable. A negative correlation indicates that, when the value of one variable increases, the value of the other decreases.

It is useful to make a distinction between correlation coefficients that are significant, and those that represent a substantive relationship between variables. For example, a correlation of .10 may be statistically significant, but is unlikely to be substantive. For this reason, we use a convention of referring to correlation coefficients as 'weak' if they range from .00 to .10, 'weak to moderate' if they range from .11 to .25, 'moderate' if they range from .26 to .40, moderate to strong if they range from .41 to .55, and 'strong' if they are greater than .55.

The number of books in students' homes, which may be considered as an indicator of general educational climate, is moderately related to achievement in each domain, ranging from .34 for combined mathematics to .37 for science. This variable is more closely related to performance than the home educational resources variable (.23 to .26), which is an indication of the availability of such materials as a desk for studying and books relevant to homework.

The correlation between number of siblings and achievement in each domain ranges from –.11 to –.14. The tabulation of number of siblings and achievement in Table 4.12 shows a non-linear relationship, suggesting that students in families with one or two siblings perform better than those in larger and in one-child families.

Mathematics self-efficacy, which was assessed through students' responses about their expected success on specified mathematical tasks, is one of the variables most strongly associated with achievement (.53). The correlation between anxiety towards mathematics and mathematics achievement is moderate and negative (–.36), indicating that students with high levels of anxiety tend to perform less well in mathematics.

Table 4.36. Linear Associations between Student Variables and Achievement in Combined Mathematics, Reading and Science

	Combined Maths			*Reading*			*Science*		
	r	*t*	*p*	*r*	*t*	*p*	*r*	*t*	*p*
Higher of parents' education	**.268**	14.07	<.001	**.237**	11.13	<.001	**.258**	12.23	<.001
Higher of parents' occupation	**.315**	14.48	<.001	**.329**	13.74	<.001	**.337**	14.54	<.001
Econom., social and cult. status	**.403**	19.71	<.001	**.406**	18.51	<.001	**.421**	21.00	<.001
Number of siblings	**-.109**	-5.33	<.001	**-.136**	-7.17	<.001	**-.123**	-6.64	<.001
Home educational resources	**.231**	12.20	<.001	**.263**	10.29	<.001	**.263**	13.60	<.001
Number of books in the home	**.335**	15.48	<.001	**.353**	14.75	<.001	**.365**	18.09	<.001
Absence	**-.185**	-8.33	<.001	**-.190**	-8.72	<.001	**-.186**	-8.91	<.001
Mathematics efficacy	**.529**	35.91	<.001						
Mathematics anxiety	**-.363**	-19.99	<.001						
Frequency of reading fiction				**.285**	14.49	<.001			
Frequency of reading e-mails/webpages				**.207**	10.21	<.001			

Note. Significant correlations are highlighted in bold. Df = 80 (number of variance strata associated with balanced repeated replication (BRR) method of variance estimation).

A number of intercorrelations for the student-level variables are in the moderate to strong range, notably between the index of economic, social and cultural status and several variables (parental occupation, parental education, books in the home, and home educational resources) (Table 4.37).

Table 4.37. Linear Associations between Student-Level Variables

	P. Occ.	P. Ed.	ESCS	Nsib	Books	Res.	Absen.	Eff.	Anx.	Fiction	E-mails
P. Occ.	1										
P. Ed.	**.488**	1									
ESCS	**.816**	**.774**	1								
Nsib	**-.148**	**-.115**	**-.158**	1							
Books	**.328**	**.344**	**.521**	-.004	1						
Res.	**.162**	**.199**	**.419**	**-.060**	**.278**	1					
Absen.	**-.089**	**-.060**	**-.112**	**.066**	**-.098**	**-.094**	1				
Eff.	**.183**	**.190**	**.282**	**-.054**	**.228**	**.214**	**-.093**	1			
Anx.	**-.111**	**-.118**	**-.147**	.012	**-.110**	**-.084**	**.096**	**-.482**	1		
Fiction	**.144**	**.129**	**.250**	-.020	**.292**	**.217**	**-.064**	**.188**	**-.073**	1	
E-mails	**.175**	**.155**	**.287**	**-.081**	**.199**	**.168**	-.002	**.185**	**-.069**	**.247**	1

Note. Significant correlations ($p \leq .05$) are marked in bold.

Variable Key

P. Occ.	Parental Occupation (SES)	Absen.	Frequency of Missing School in past Two Weeks
P. Ed.	Parental Education	Eff.	Mathematics Self-Efficacy
ESCS	Economic, Social and Cultural Status	Anx.	Anxiety About Mathematics
Books	Number of Books in the Home	Fiction	Frequency of Reading Fiction
Nsib	Number of Siblings	E-mails	Frequency of Reading E-mails and Webpages
Res.	Educational Resources in the Home		

This is not surprising since the ESCS measure was constructed from these aspects of home background. Other variables that are positively correlated are those between parental occupation and parental education (.49), between parental occupation and number of books in the home (.33), and between parental education and number of books in the home (.34). There is also a moderate to strong negative correlation between mathematics self-efficacy and anxiety about mathematics (–.48).

School-Level Variables

The school-level variables that correlate most strongly with student achievement include the economic, social and cultural status (ESCS) of the school (correlations with achievement in the three assessment domains range from .36 to .40) (Table 4.38). Correlations between the percent of students entitled to Junior Certificate Examination fee waivers and achievement in the three domains are in the moderate range (–.29 to –.32). The relationship between achievement and the total length (in minutes) of overall instructional time per week is weak to moderate, but positive, ranging from .15 to .17. The percent of 15-year-old females shows weak to moderate positive relationship with achievement in the case of reading literacy (.10); in the other two domains, it is weak and negative, and, in the case of science it is not statistically significant. There is also a weak to moderate negative relationship (ranging from –.19 to –.22) between the ratio of computers to students and their achievement, but this relationship appears to be mediated by the SES composition of the school. The associations between disciplinary climate in mathematics lessons and achievement are also in the weak to moderate range (.19 to .23).

Table 4.38. Linear Associations between School Variables and Achievement in Combined Mathematics, Reading and Science

	Maths (Comb.)			*Reading*			*Science*		
	r	*t*	*p*	*r*	*t*	*p*	*r*	*t*	*p*
Junior Cert. fee waiver	**-.308**	-9.48	<.001	**-.287**	-9.80	<.001	**-.317**	-10.72	<.001
Econom. social and cult. status**	**.363**	14.85	<.001	**.395**	12.77	<.001	**.363**	15.74	<.001
Percent of 15-year-old females	**-.064**	-2.18	.032	**.103**	3.28	.002	-.015	-0.51	.608
Ratio of computers to students*	**-.185**	-6.21	<.001	**-.221**	-6.92	<.001	**-.194**	-6.77	<.001
Tot. min. of instructional time per week*	**.170**	7.27	<.001	**.165**	7.26	<.001	**.145**	6.48	<.001
Disciplinary climate in maths classes**	**.211**	9.71	<.001	**.227**	10.02	<.001	**.188**	8.30	<.001

Note. Significant correlations are highlighted in bold. Df = 80 (number of variance strata associated with balanced repeated replication (BRR) method of variance estimation).
* Variables have missing values over 10%.
** Variable derived from the Student Questionnaire.

As with the student-level variables, some of the school-level variables are inter-related. Table 4.39 shows the inter-correlations for all continuous school-level variables in Table 4.38. The percent of students entitled to a fee waiver for the Junior Certificate is strongly correlated with the school-level economic, social and cultural status (–.81). It also has a moderate/strong correlation with the computer-student ratio (.50). School-level economic, social and cultural status is a moderately strongly correlated with the computer-student ratio (–.38).

Table 4.39. Linear Associations between School-Level Variables

	JCE Fee Waiver	Sch. ESCS	PCFemale	Ratio of comp.	Total Instr.	Disc. Clim.
JCE Fee Waiver	1					
Sch. ESCS	**-.808**					
PCFemale	.013	**.004**	1			
Ratio of comp.	**.495**	**-.381**	-.022	1		
Total Instr.	**-.098**	**.082**	.011	**-.083**	1	
Disc. Clim.	**-.061**	**.078**	**.200**	**-.060**	.004	1

Note. Significant correlations ($p \leq .05$) are marked in bold.

Variable Key

JCE Fee Waiver	Junior Certificate Examination fee waiver
Sch. ESCS	School-Level Economic, Social and Cultural Status
PCFemale	Percent of 15-year-old females
Ratio of comp.	Computer-Student Ratio
Total instr.	Total minutes of instructional time per week
Disc. Clim.	Disciplinary Climate in Maths Classes

CONCLUSION

Males in Ireland significantly outperformed females on the combined mathematics scale (by about one-sixth of a standard deviation). The largest difference was found on the Space & Shape subscale (almost one-third of a standard deviation). Irish females significantly outperformed males on reading literacy (by about one-third of a standard deviation). There was no significant difference in Irish males' and females' performance on science. The pattern of gender differences observed across the three domains is consistent with most other participating countries, and with gender differences in Ireland in PISA 2000.

The population of Irish 15-year-olds is comparatively homogeneous; 95.0% of students and their parents were born in Ireland. Country of birth was not related to performance outcomes, in contrast to the findings for many OECD countries. Large differences in achievement, in the order of three-quarters of a standard deviation, between Irish settled and Traveller students were found (although just 0.7% of participating students identified themselves as Travellers).

Students from higher-SES backgrounds perform significantly better than students from lower-SES backgrounds. The magnitude of the advantage associated with high SES is about the same across all three domains.

The 15% of Irish students from lone-parent families, scored significantly lower than students from other types of family in all three domains. The difference in achievement scores between students from lone- and dual-parent families (around two-fifths of a standard deviation in all three domains), is the third highest across all participating countries. A high number of siblings was also negatively related to student performance. Students with three or more siblings had lower mean scores than students with one sibling.

Home educational resources were found to be related to achievement. For example, students with high levels of educational resources (a quiet place to study, a desk for study, and books to help with schoolwork) scored about half a standard deviation higher than students with access to none or one of these resources, and students with high numbers of books in the home (500+) scored over two standard deviations higher on the reading literacy scale than students with very few or no books (none to 10 books).

About one-fifth of students (20.2%) were judged to be at risk of leaving school prior to completion of the Leaving Certificate, which is consistent with national rates of early school leaving. These students score significantly lower in all three domains (by around one standard deviation) than students who intended to complete a Leaving Certificate programme. The percentage of students at risk in 2003 is somewhat higher than the percent in 2000 (14.0%).

Students' current grade level was also related to achievement. Students in Second year had mean scores on all three domains that were about one standard deviation lower than students in Third year (the modal grade level).

Student absenteeism (as indicated by the number of days missed in the fortnight prior to the PISA assessment) was also associated with achievement, with a difference of about two-thirds of a standard deviation in all three domains favouring students missing no days compared to students missing three or more days. Data were not collected on the reasons for these absences.

Overall time spent on homework and study each week showed positive associations with achievement, with a difference of about one-third of a standard deviation between the achievements of students in the upper and lower thirds of the distribution. There are no significant differences in the mean scores of students in the medium and high groups of this variable.

Large differences in achievement (amounting to about one standard deviation) were found in all three PISA domains between students taking the corresponding Junior Certificate subject at Higher and Ordinary levels. The mean score difference on PISA mathematics and PISA reading literacy was around one standard deviation between students at Ordinary and Foundation levels. In the case of science (which is an optional subject at Junior Certificate level), the mean score difference between students taking the subject at Ordinary level and those not taking it is around nine score points (around one-tenth of standard deviation) and is not significant. In both reading and mathematics, the mean scores of Foundation level students are at the lowest proficiency level in the corresponding PISA domains (Level 1).

The majority (78.0%) reported using a calculator in the PISA assessment. This is higher than the percentage in 2000 (24.2%). Students using a calculator in 2003 had a significantly higher mean score than those that did not (with a score difference in the order of one-quarter of a standard deviation).

Students' self-reports about their perceived efficacy at completing various mathematics tasks show a particularly strong association with achievement: the score difference in mathematics for students reporting low and high self-efficacy is about one and one-quarter standard deviations. Because the self-efficacy measure is based on students' ratings of their own confidence at completing specific mathematics tasks it itself may be indicative of current mathematics achievement, or indeed, experience of taking the PISA mathematics test, which preceded administration of the Student Questionnaire.

Students' reports of levels of anxiety in dealing with mathematics in testing, class work and study contexts also revealed differences, although in the opposite direction: higher anxiety was associated with lower performance. Students reporting high levels of anxiety had a mean mathematics score that is about four-fifths of a standard deviation lower than the mean score of students with low levels. Irish females reported significantly higher levels of anxiety than males (the difference is about one-quarter of a standard deviation), a pattern that was found in many other countries.

Significant achievement differences between schools appear to relate mainly to school type, structure and school socioeconomic status. For example, there are differences between the vocational, secondary and community/comprehensive sectors, with the largest differences (around half a standard deviation in the three assessment domains) between vocational and secondary schools. These are similar to the achievement differentials in PISA 2000. In each domain, the achievements of students in schools that were designated as disadvantaged are about two-fifths of a standard deviation lower than the achievements of students in non-designated schools. The differences are somewhat smaller than those observed in PISA 2000.

A large difference of around three-quarters of a standard deviation was observed between students in schools in Ireland categorised as high and low in economic, social and cultural status (ESCS). The ESCS composite variable combines aspects of parental education, parental occupation, and home possessions. As such, it is a rather complex variable which may not be easily applied in policy development. Another variable relating to school-level socioeconomic status is the percentage of students entitled to a fee waiver for the Junior Certificate Examination. This showed a substantial association with achievement. Students in schools with a low rate of fee waiver scored seven-tenths to four-fifths of a standard deviation higher (depending on the domain) than those in schools with a high rate of fee waiver.

When the achievements of students in schools with different gender compositions are compared, an interesting pattern emerges. Students in all boys schools score significantly higher in mathematics than students in all girls and mixed sex schools. While students in all girls schools score significantly higher in reading than students in all boys schools. Differences in performance across schools of different sex composition may be related to socioeconomic status.

Students in schools with a low (poor) disciplinary climate in mathematics classes (as reported by students and aggregated to the level of the school) achieved mean scores that are about half a standard deviation lower than students in schools reporting a high (good) climate. Disciplinary climate seems to operate independently of SES compositions, since its correlations with school-level economic, social and cultural status and with percent of students entitled to fee waiver for the Junior Certificate are weak (less than .08).

In some countries participating in PISA, the composition of the student population is culturally heterogenous, and in several of these, there are notable performance gaps between native and first-generation students. The Irish population is still comparatively homogenous, with 95% of Irish students indicating that both they and their parents were born in Ireland. Given that foreign-born students are increasingly likely to attend Irish schools (Eurydice, 2004), the maintenance of homogeneity in achievement outcomes merits careful monitoring.

5

Explanatory Models of Irish Performance in Mathematics, Reading and Science in PISA 2003

In Chapter 4, relationships between a variety of background variables and performance in PISA were presented for Irish students. Since many of these variables are interrelated, caution is needed in considering the separate association of each with achievement. This is because an observed relationship between a variable and student performance may be because both are related to a third variable (and possibly others). This chapter looks at the conditional relationship of explanatory variables with student performance when a range of others is considered simultaneously, reducing the possibility of spurious associations that may arise when only one variable at a time is considered.

Furthermore, since achievement scores show a form of dependency among students in each school, known as clustering, multilevel regression models which enable the total variation in student performance to be split into between- and within-cluster components are used. These also allow variables to be fitted at different levels (school, class and student) (Goldstein, 1987, 1995; Longford, 1993). This modelling requires the selection of key variables and the estimation of how achievement differs as each explanatory variable changes in value, adjusting for the other variables included in the model. Statistical measures calculated during the process of model-building allow the evaluation of its adequacy in explaining the observed patterns of student achievement.

The final models for each of the three assessment domains, along with interpretations, are presented following a description of procedures used to develop the models. Readers with a less technical interest in the analyses might benefit from skimming the description of procedures, and proceeding directly to the final models and their interpretations.

BETWEEN- AND WITHIN-SCHOOL VARIANCE COMPONENTS OF ACHIEVEMENT

As described in Chapter 1, the design of PISA 2003 involved the random selection of 35 students, where available, from each of 150 schools.[36] The sampling induces an intra-cluster correlation (ICC) between student scores in each school, since students in a school tend to be more similar to one another than students sampled at random across schools. Total variation in achievement can be partitioned into a between-cluster component, i.e. school variation, and a within-cluster component, that of class plus student variation. The ICC corresponds to the between-cluster variance component expressed as a proportion of the total variation. The latter is often converted to a percentage for presentation purposes.

Table 5.1 shows the percentage of total variation in student scores in combined mathematics, reading and science that lies between schools for all countries participating in PISA 2003. Postlethwaite (1995) has suggested that large between-school variance components are suggestive of more heterogeneous school systems. The ICC for Ireland for combined mathematics (16.7%) is low compared to most of the countries in the table. The

[36] This is the standard PISA 2003 sample design. Some countries, such as Germany and Canada, had larger sample sizes. Ireland sampled 154 rather than 150 schools to account for the fact that some Irish schools are small in size, containing fewer than the required 35 students aged 15.

value for Ireland is close to those for Canada, New Zealand, Spain, and partner country Macao-China.

Table 5.1. Percentages of Total Variance in Achievement in Combined Mathematics, Reading and Science that Lie Between Schools – OECD and Partner Countries

	Mathematics	*Reading*	*Science*
Iceland	4.0	4.0	4.1
Finland	4.9	3.9	4.4
Norway	6.9	7.8	7.9
Sweden	10.8	9.6	9.2
Poland	12.7	14.7	14.2
Denmark	13.7	18.1	12.6
Canada	16.3	14.8	14.8
Ireland	**16.7**	**22.5**	**16.2**
New Zealand	18.6	17.8	18.2
Macao-China	19.1	23.9	17.5
Spain	19.7	19.2	17.2
Australia	21.4	21.3	20.3
Latvia	22.7	20.2	20.6
United States	25.8	24.1	22.0
Russian Federation	30.2	23.5	21.1
Luxembourg	32.2	27.5	29.2
Portugal	33.8	37.8	31.0
Switzerland	34.2	30.1	30.0
Serbia	35.5	34.6	28.9
Greece	36.6	34.9	26.9
Thailand	37.3	33.7	32.0
Mexico	38.3	35.2	27.7
Korea	42.1	35.9	38.3
Tunisia	42.1	33.2	33.7
Slovak Republic	43.0	41.1	42.6
Uruguay	44.4	36.2	33.6
Indonesia	44.7	36.5	37.2
Liechtenstein	44.8	45.1	42.0
Brazil	45.1	n/a	n/a
France	45.9	45.4	47.5
Hong Kong-China	46.7	42.0	45.4
Czech Republic	47.7	41.4	39.0
Belgium	51.0	50.9	45.5
Italy	52.3	48.7	48.0
Germany	52.4	52.7	51.1
Austria	52.7	55.9	53.5
Japan	53.3	44.6	46.3
Turkey	54.9	52.7	52.8
Hungary	57.9	53.0	50.3
Netherlands	58.2	53.7	54.7
OECD Average	**32.7**	**31.4**	**29.9**

Note. OECD countries are in regular font, partner countries are in *italics* . Countries are ordered in descending order of the ICC for mathematics. Data for Brazil are not available for reading and science.

In Scandinavian countries, the ICC is very low – less than 10% in Iceland, Finland and Norway. In contrast, the ICC exceeds 50% in eight participating countries (Belgium, Italy, Germany, Austria, Japan, Turkey, Hungary, and the Netherlands), indicating greater heterogeneity amongst schools and hence, more selective school systems. The ICC for Ireland for reading literacy (22.5%) is a little higher than combined mathematics, but nonetheless lower than for most participating countries, while in science, it is almost identical to combined mathematics (16.2%).

PROCEDURES USED FOR MULTILEVEL MODELLING

The type of multilevel models used were *hierarchical linear models*. These are linear regression models with random components at the cluster and individual level. The most basic version simplifies the variation in the intercept that occurs from school to school (the clusters) by fitting a random effect which 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 may also be made to vary across clusters (i.e., the effect of student-level variables may differ across schools) by including random effects for their parameters, known as random coefficients. This will suggest a range of values that the parameter estimate, e.g., the gender difference, takes over the population of clusters (schools), with the fitted Normal distribution for the random effect. The observations are assumed to be independent by conditioning on the random effects, namely the random intercept and any random coefficients fitted in the model. The NLME library of Pinheiro and Bates[37] (2000) implemented in the R statistical package was used for fitting the multilevel models. This was extended to deal with the five plausible values corresponding to each student's score (see Chapter 1 for a description of plausible values).

The explanatory variables chosen for evaluation during the process of model-building either showed statistically significant associations with achievement (described in Chapter 4), or were of sufficient policy or theoretical interest to merit inclusion. Another consideration was the amount of missing data for a variable – those with less than 5% missing data (or 'missingness') were preferred over similar variables with higher levels. When some of these variables were highly correlated or theoretically linked, a composite variable was constructed or the one with the largest association with student achievement was chosen. For example, home educational resources is constructed from possession of each of the following: a desk, a place to study, textbooks (high group), missing one of the items (medium group), or missing two or more items (low group). Avoiding variables that are highly inter-related reduces problems caused by multi-collinearity (Hutcheson & Sofroniou, 1999) and facilitates interpretation of the chosen model.

As recommended by Aitkin, Francis and Hinde (in press), the models are unweighted, involving an evaluation of the explicit stratifying variable (school size) during the model-building. When not already centred at zero, the continuous explanatory variables (variates) were centred by subtracting their weighted mean, i.e., socioeconomic status (mean = 48.3), school disciplinary climate in mathematics (mean = 0.27) and percent fee waiver for the Junior Certificate Examination (mean = 25.8). This improves numerical estimation by the statistical software while giving a convenient interpretation to the intercept; i.e., the expected score for a student who was at the mean value of the original continuous explanatory variables. The categorical explanatory variables (factors) were internally dummy coded by

[37] 'NLME' stands for non-linear mixed effects.

the software used and had their reference category set to the same group as used in Chapter 4 to enable straightforward comparison.[38]

An initial examination of the curvilinearity[39] of the relationship between each variate and one of the five plausible values of the achievement score was carried out by means of ordinary least-squares (OLS) regressions, graphical displays of residuals and scatterplot smoothing (Cook & Weisberg, 1999). Any curvilinear terms, e.g., squared or logarithmic functions of variates, were subsequently evaluated using hierarchical linear models with all five plausible values of achievement.

Models were estimated by full maximum likelihood, which allows deviance tests of the significance of both fixed and random effects. The deviance is a measure of the goodness-of-fit of a model. Categorical variables and variance components were evaluated using omnibus tests[40] of deviance differences for models fitted with and without the factor, referred to a χ^2 distribution, with the degrees of freedom set to the difference in the number of fitted terms (McCullagh & Nelder, 1989). Following the practice of the HLM6 software package (Raudenbush et al., 2004), averages in statistics tested by χ^2 across the plausible value datasets were used, with degrees of freedom equal to their complete data values. This method was used for testing deviance differences. The degrees of freedom for tests of deviance differences were set to the difference in the number of terms between nested models. The tests for single parameters (e.g., for variates and factors with two categories) used the adjusted *t* tests of Little and Schenker (1995). In addition, the formulae provided by these authors for combining the parameter estimates and standard errors calculated from each plausible value dataset were used to generate the final estimates and standard errors.

The level of missingness of each explanatory variable was below 5%. Therefore the non-missing indicator variable method used in the national report for PISA 2000 (see Shiel et al., 2001, page 102) was not required (see also Lindsey & Lindsey, 2001 for more detail on the non-missing indicator method). Rather, listwise deletion[41] was applied to the missing cases in the present analysis.

Separate models of achievement for each explanatory variable with random intercepts were initially fitted to evaluate whether, when variables were subsequently entered simultaneously, the parameter estimates changed in any substantive way. This would suggest that the explanatory variables were sensitive to the other variables present and related to each other in a complex fashion.

All the candidate variables (both at the student and school level) were then entered simultaneously as main effects into a single model. Non-significant variables were removed using a manual backwards elimination strategy using a criterion of $p \leq .05$ for each test. The significance of variables remaining in the model was re-evaluated each time one was removed. Once a model with all significant main effects terms was obtained, three omnibus tests of all two-way interactions amongst student-level variables, all cross-level interactions, and tests for all school-level variables were carried out, keeping the error rate down

[38] A *dummy variable* is a numerical variable used in regression analysis to represent subgroups. In the models presented in this chapter, dummy variables with values 0, 1 are used, where a student is given a value of 0 if he/she is not in a group (e.g., not in a school designated as disadvantaged) or a 1 if he/she is in the group (e.g., in a school designated as disadvantaged). Dummy variables are useful because they enable one to use a single regression equation to represent multiple groups.

[39] A relationship is said to be *curvilinear* if the amount of change in the response variable is not constant for each unit change in the explanatory variable. This is evident in scatterplots of the variables in question, which show a curved rather than straight line; the significance of the curvilinearity of the relationship can be tested formally using regression techniques.

[40] In this context, an *omnibus test* assesses the improvement in model fit by adding all levels of the categorical variable simultaneously.

[41] 'Listwise deletion', a standard procedure in many statistical software packages, refers to cases being dropped from the dataset if missing on any of the set of explanatory variables.

compared to the large number of individual tests corresponding to these. Statistical significance in one of these omnibus tests was then followed by omnibus tests of each corresponding variable in interaction with the remaining ones defined by the original test. For example, if the test for adding all student-level interactions was significant, an omnibus test of all the gender interactions with the other student-level variables would be carried out, followed by a similar omnibus test of the SES interactions, etc.. Finally, if two of these single-variable omnibus tests proved significant, the interaction term of the two variables would be evaluated.

Following selection of the main effects and interaction terms, the curvilinearity of the variates was evaluated by testing the significance of adding the corresponding orthogonal polynomial terms[42] for a variable. An exception was the index of books in the home, which was fitted more parsimoniously at the outset with a logarithmic function.

Finally, all the student-level variables were evaluated to see whether their effects could be considered constant across schools by the addition of random coefficients for each term, one at a time, using tests of deviance changes. Factors with more than two categories had all the random coefficients for the corresponding set of parameter estimates added. Inset 5.1 gives several points to assist in interpreting the tables of in this chapter.

Inset 5.1. Interpreting the Tables of Multilevel Models

The following points may be borne in mind when interpreting the tables in this chapter.

- The estimates in all tables are unweighted. When variables are added to the null (empty) model separately (e.g., Table 5.2), because the estimates are unweighted, they do not correspond exactly to the estimates provided in Chapter 4. The listwise deletion of cases with missing values (i.e., removal of all cases with one or more missing values on the explanatory variables) implemented in the *R* software will also lead to a difference in the estimates.
- Continuous variables, such as socioeconomic status, have been centred around the grand mean. This results in an intercept corresponding to the predicted achievement of a student with a mean value on each continuous variable.
- For categorical variables (such as school type), the reference category is given alongside the label for the category corresponding to the parameter estimate.
- Where an interaction term is included, the formal significance test for the main effect is omitted since main effects cannot be sensibly evaluated in the presence of interactions involving them. This also applies to the linear term in a quadratic fit; i.e., only the squared term can be evaluated when ordinary polynomials are used.
- In describing the tables, the parameter estimates are translated into units of standard deviation. It is useful to bear in mind, assuming that student achievement is Normally distributed, that one standard deviation above and below the mean accounts for 68% of students' scores, and two standard deviations above and below the mean account for roughly 95% of scores. The standard deviations for Ireland for each of the three achievement scales are as follows: mathematics: 85.3; reading literacy: 86.5; and science: 93.0.
- When interpreting the variance components for the final models (such as that for mathematics shown in Table 5.4), it is useful to note that the square root of the variance is the standard deviation. For example, the variance component for the intercept of the final model of mathematics is 172.169; the standard deviation is 13.121. For random intercepts, the standard deviation can be used to calculate the proportion of schools expected within a given range of mean scores.

[42] *Orthogonal polynomials* are uncorrelated versions of conventional polynomials, e.g. quadratic ($x + x^2$) and cubic ($x + x^2 + x^3$).

Characteristics of the Analysed Subsample

Of the 3880 cases available for analysis, 3501 were complete for the variables used to model combined mathematics (90.2% of cases), 3518 for reading literacy (90.7% of cases), and 3478 for science (89.6% of cases). The typical cluster size used in R-squared calculations was based on the mean number of 15-year-old students enrolled in a school, with a value of 82.1.[43]

MODELS OF COMBINED MATHEMATICS

Development of the Model

Table 5.2 gives the parameter estimates, standard errors and significance tests of each student-level variable fitted separately, while Table 5.3 gives the corresponding values for the school-level variables. The natural logarithm (referred to as log) of the index of books in the home was used. Factors with more than two categories have several parameter estimates and these are evaluated by a single omnibus test for their inclusion. All the separate models involving the candidate student- and school-level variables provide statistically significant improvements over the null model with just a random intercept. Some examples of interpretations of the parameter estimates are as follows: there is a 17.9-point (one-fifth of a standard deviation) deficit in scores for females compared to males; students attending community or comprehensive schools score on average 19.0 points (just over one-fifth of a standard deviation) lower than students in secondary schools, while students in vocational schools score 39.9 points (close to half a standard deviation) lower than students in secondary schools.

Table 5.2. Achievement in Mathematics: All Student-Level Variables Tested as Separate Models by Addition to the Null Random Intercept Model

	Parameter	*SE*	*Test Statistic*	*df*	*p–value*
Gender: female–male	–17.875	3.406	t = –5.429	25833	<.001
Socioeconomic status	1.359	0.095	t =14.332	138	<.001
Lone parent: dual–lone	26.841	3.812	t = 7.041	6265	<.001
Number of siblings			Ddiff = 56.249	4	<.001
none–one	–20.574	5.276			
two–one	–10.022	4.025			
three–one	–16.557	4.451			
four or more–one	–29.626	4.582			
Log books index	47.132	2.872	t = 16.412	820	<.001
Home educational resources			Ddiff = 75.649	2	<.001
low–high	–30.146	4.107			
medium–high	–17.677	3.187			
Absence			Ddiff = 99.062	2	<.001
1 or 2 days–no days	–17.066	3.333			
3 days or more–no days	–42.439	5.362			
Grade level			Ddiff = 268.033	3	<.001
Grade 8–Grade 9	–69.739	8.475			
Grade 10–Grade 9	41.241	4.579			
Grade 11–Grade 9	30.863	3.808			

Note: Ddiff = deviance difference, tested using the χ^2 distribution. Categorical variables with two or more levels are in italics.

[43] This figure is the average of the number of 15-year-olds enrolled in the schools on the PISA 2003 sampling frame for Ireland.

Table 5.3. Achievement in Mathematics: All School-Level Variables Tested as Separate Models by Addition to the Null Random Intercept Model

	Parameter	*SE*	*Test Statistic*	*df*	*p–value*
Size (stratum)			Ddiff = 9.888	2	.007
small–large	–33.066	15.016			
medium–large	–16.379	6.639			
Sector			Ddiff = 34.867	2	<.001
comm/comp–secondary	–18.989	7.190			
vocational–secondary	–39.893	6.824			
Disciplinary climate	16.225	7.678	t = 2.113	143	.036
Percent JCE fee waiver	–1.591	0.154	–10.351	143	<.001

Note: Ddiff = deviance difference, tested using the χ^2 distribution. JCE = Junior Certificate Examination. Categorical variables with two or more levels are in italics.

After this initial examination, all the variables were fitted as main effects entered simultaneously into the model. The non-significant variables, school size (stratum) and school sector were removed in sequence with the significance of the remaining terms re-evaluated each time one was eliminated. The omnibus tests of the student-, cross- and school-level two-way interactions, followed by the follow-up tests, suggested the terms corresponding to a log books index × absence interaction were required. No significant polynomial terms were found, indicating that a linear model for the variates apart from log books index is sufficient. An examination of the random coefficients for the student-level variables, evaluated one term at a time by addition to this model (or set of terms at a time, e.g., in the case of factors) failed to provide any significant improvement. This suggests that the student-level parameter estimates were constant across the clusters (schools).

The Final Model

Table 5.4 presents the final model of combined mathematics, its parameter estimates, standard errors, and significance tests.

Significance tests are not given when a variable appears in a higher-order interaction (McCullagh & Nelder, 1989); for example, log books index has an interaction with absence and therefore both are present as main effects and only the interaction term is evaluated in the table. Categorical variables with more than two levels are tested in a single omnibus evaluation of their removal and contain more than one parameter estimate. For example, the four terms for number of siblings are removed in one go.

Comparing Table 5.4 to the separately fitted models for each variable (Tables 5.2 and 5.3), the gender difference has shown a small increase and the deficits corresponding to increased absence from school are now prominent for both degrees of absence. School stratum (size) and school sector, though significant when evaluated separately in relation to achievement, both dropped out of the model. This indicates that the variation in scores corresponding to them has been explained by other terms retained in the final model.

Table 5.4. Final Model of Achievement in Combined Mathematics

	Parameter	*SE*	*Test Statistic*	*df*	*p–value*
Intercept	473.044	6.575			
Student–Level Main Effects					
Gender: female–male	–24.223	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) × Absence			Ddiff = 7.355	2	.025
Log (books index) × 1 or 2	–12.615	5.737			
Log (books index) × 3 or more	–15.069	8.534			
Variance Components					
Intercept	172.169				
Residual	4722.511				

Note: Ddiff = deviance difference, tested using the χ^2 distribution. JCE = Junior Certificate Examination. Variables dropped during model-building (in sequence): School size (stratum) and school sector. Categorical variables with two or more levels are in italics.

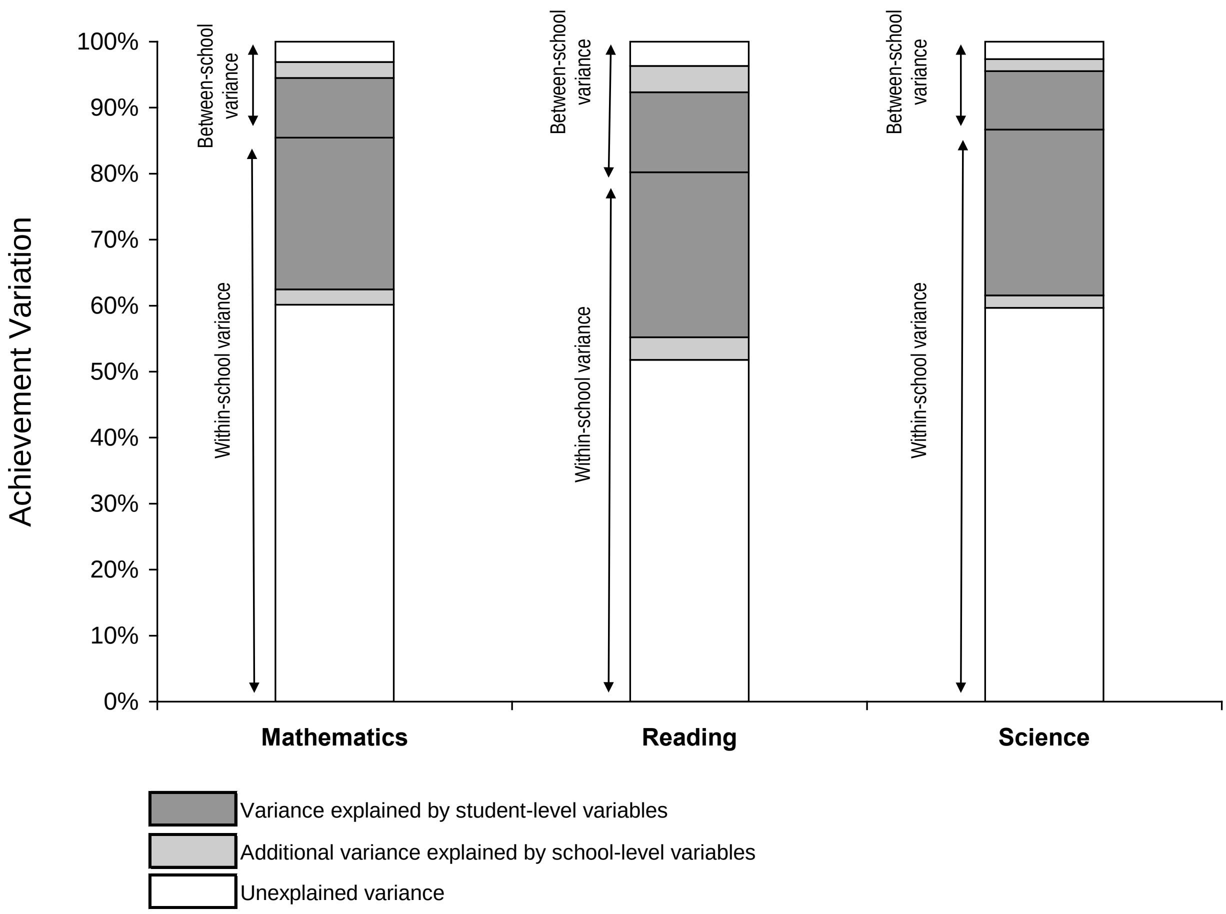
The ICC of the null random intercept model suggested that 14.6% of the total variation in scores was at the cluster (school) level.[44] The formula used to calculate the explained variation at each level is from Snijders and Bosker (1999) and is given in Inset 5.2. The sub-model which fits the student-level variables and interactions alone gives explained variances of 62.3% at the cluster-level and 26.9% at the individual (class and student) level. Including the school-level variables gives improvements of an additional 16.5% and 2.7%, respectively. So the final model explains 78.8% of the between-cluster variation and 29.6% of the variation within clusters.

[44] This value differs somewhat to that observed for all participating students shown in Table 5.1 as not all had complete data for all variables in the model and because full maximum likelihood estimation was used for the models.

Figure 5.1 shows the total variation in achievement in combined mathematics (as well as the final models of reading literacy and science) divided into variance explained by the model (both at the student level and at the level of the school) and variance that is left unexplained.

> ***Inset 5.2. Calculation of the Proportion of Explained Variance in Achievement***
>
> The method used to calculate the proportion of variance in achievement at the cluster level (Snijders & Bosker, 1999) requires one to use a representative value for the size of the clusters. The mean enrolment size (82.1) of all schools in the PISA sampling frame (the desired population) was used for the representative cluster size (CS). The formulae used were:
>
> Level 1 R^2 = *1 – (Var L1F + VarL2F)/(VarL1N + VarL2N)*
> Level 2 R^2 = *1 – (Var L1F/CS + VarL2F)/(VarL1N/CS + VarL2N)*
> Where
>
> VarL1F = Level 1 variance of fitted model
> VarL2F = Level 2 variance of fitted model
> VarL1N = Level 1 null model variance
> VarL2N = Level 2 null model variance

Figure 5.1. *Variance Explained by Student-Level Variables, Additional Variance Explained by School-Level Variables, and Unexplained Variance: Final Models of Combined Mathematics, Reading Literacy and Science*



Contributions of the Explanatory Variables to Fitted Scores in Combined Mathematics

The chosen reference categories for factors and mean-centred variates (socioeconomic status, school disciplinary climate in mathematics classes, and percent entitled to fee waiver for the Junior Certificate Examination) give an intercept of 473.0. This corresponds to a hypothetical student who is male, of average SES, in a lone-parent family, with one sibling, with zero to 10 books in the home (log of book index 1 = 0), with high home educational resources, no absences from school in the fortnight prior to the PISA assessment, in Grade 9 (Third Year), in a school with average disciplinary climate in mathematics classes, and an average rate of fee waiver for the Junior Certificate Examination. Fitted values for other combinations of student and school variables can be calculated by inserting the contribution of each to the fitted score and adding these values. The parameter estimates for continuous explanatory variables are first multiplied by the chosen example value of the variate (first subtracting the mean of the variable if it was centred around its mean during model fitting). In contrast, the values for categorical variables can be read directly from the parameter estimates in the table. Thus, for example, the gender difference associated with the final model of mathematics (Table 5.4), adjusting for the other variables in the model, corresponds to a deficit of 24.2 points (about three-tenths of a standard deviation) for females compared to males. Students from dual-parent families tend to score about 15.7 points (one-sixth of a standard deviation) higher than students from lone-parent families. A student who is in Grade 8 (Second Year) scores on average 59.2 points (two-fifths of a standard deviation) lower than a student in Grade 9 (Third Year), while students in Grades 10 and 11 (Fourth Year and Fifth Year) score higher by values of 33.9 and 35.7 respectively.

A continuous variable which appears linearly as only a main effect can be interpreted as the change in the achievement score corresponding to a one-unit increase in the explanatory variable. Since this value will depend on the scale on which the variate was measured, we fit example values for the explanatory variable at the mean of the first, middle and upper thirds of population on that variable and present the contributions to the fitted scores in Tables 5.5 to 5.7.

Table 5.5 presents the contributions to fitted scores for values of SES. First the raw SES values are given, then the corresponding centred values, followed by the score contributions from the model. Compared to a student at the mean of the high group of the SES variable, students at the means of the medium and low groups are expected to score 14.9 and 28.6 points (about one-sixth and one-third of a standard deviation) lower respectively.

Table 5.5. Contribution to Fitted Scores in Combined Mathematics for Example Values of Student Socioeconomic Status (SES)

	SES		
	Low	*Medium*	*High*
Raw SES Score	30.54	47.88	66.59
Centred SES Score	–17.95	–0.61	18.10
Contribution to Student Score	–14.26	–0.48	14.37

Note. SES values are centred by subtracting the weighted SES mean (48.34) before the fitted values are calculated.

Table 5.6 gives the score contributions for values of disciplinary climate in mathematics classes. Students in schools at the mean of the medium disciplinary climate variable score 25.5 points (three-tenths of a standard deviation) lower than their counterparts at the mean of the high level, while the corresponding value for a student at the mean of the low level is 52.1 points (three-fifths of a standard deviation) lower than a student at the mean of the high level.

Table 5.6. Contribution to Fitted Scores in Combined Mathematics for Example Values of Disciplinary Climate in Mathematics Classes

	Disciplinary Climate		
	Low	*Medium*	*High*
Raw Disciplinary Climate Score	–1.09	0.29	1.62
Centred Disciplinary Climate Score	–1.35	0.03	1.36
Contribution to Student Score	–26.03	0.56	26.07

Note. Disciplinary climate values are centred by subtracting the weighted disciplinary climate mean (0.2651) before the fitted values are calculated.

The contributions to fitted scores for percent fee waiver for the Junior Certificate Examination are given in Table 5.7. Compared to students in schools at the mean of high fee waiver rates, those at the mean of the medium level score 16.6 points (one-sixth of a standard deviation) higher. Students at the mean of the low fee waiver level score 26.5 points (one-third of a standard deviation) higher than a student at the mean of the high level. Note that this contribution occurs after adjusting for student-level SES.

Table 5.7. Contribution to Fitted Scores in Combined Mathematics for Example Values of Fee Waiver for the Junior Certificate Examination

	JCE Fee Waiver		
	Low	*Medium*	*High*
Raw JCE Fee Waiver Score	10.67	23.07	43.80
Centred JCE Fee Waiver Score	–15.17	–2.77	17.96
Contribution to Student Score	12.14	2.21	–14.37

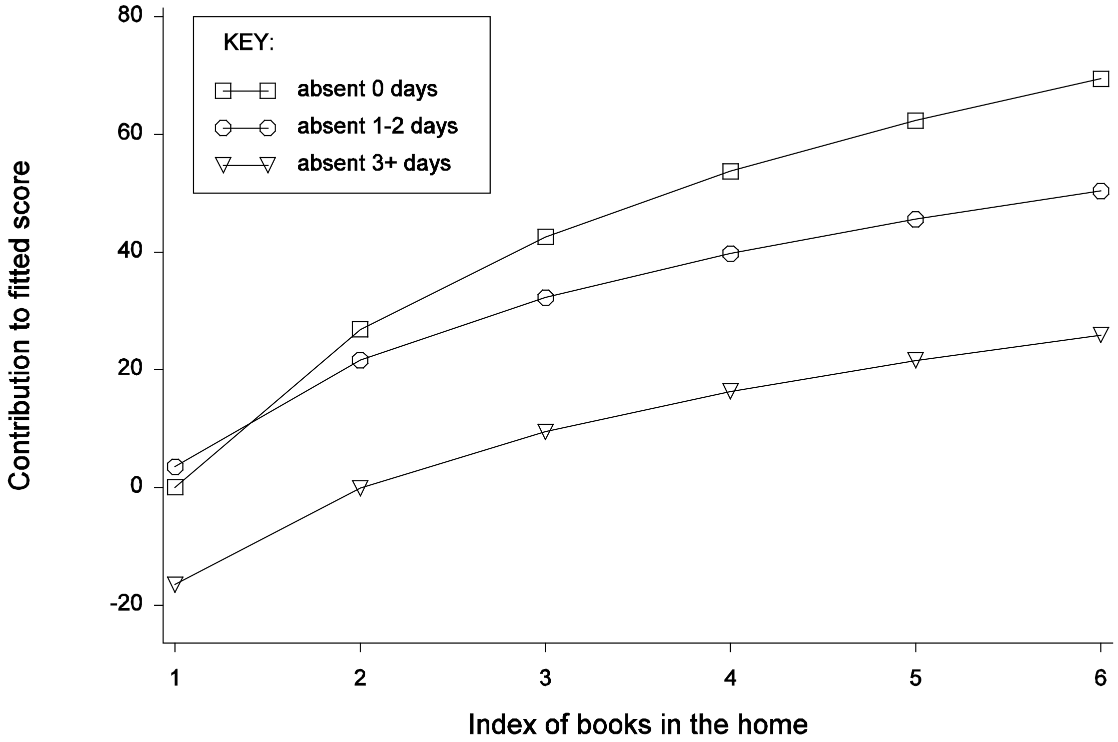
Note. JCE fee waiver values are centred by subtracting the weighted JCE fee waiver mean (25.84) before the fitted values are calculated.

The log books index variable is involved in an interaction with the parameters for frequency of absenteeism. Here, calculations using several parameter estimates are combined to produce a contribution to the fitted score. These are given for each of the six levels of the book index in Table 5.8 and are plotted in Figure 5.2. The impact of books in the home levels off as the upper value is approached, with a steeper gradient displayed for the curve corresponding to no days' absence. It also suggests that there is little difference between the none and 1 to 2 days absence score contributions for students in homes with no books, but that the large deficit for 3 or more days remains even for that group of students (at least 17 score points). The large effect corresponding to the books index is confirmed for all students, e.g., 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 (half a standard deviation) for those absent 3 or more days.

Table 5.8. Contribution to Fitted Scores in Combined Mathematics for Books in the Home by Absence from School

	Book Index					
Absence	*None to 10 (1)*	*11 to 25 (2)*	*26 to 100 (3)*	*101 to 200 (4)*	*201 to 500 (5)*	*500+ (6)*
None	0.00	26.86	42.57	53.71	62.36	69.42
1 or 2	3.53	21.64	32.24	39.76	45.59	50.35
3 or more	–16.53	–0.12	9.48	16.29	21.57	25.89

Figure 5.2. *Contribution to Fitted Scores in Combined Mathematics for Books in the Home by Absence from School*



Note. Actual values corresponding to the index of books are as follows: 1 = none to 10; 2 = 11 to 25; 3 = 26 to 100; 4 = 101 to 200; 5 = 201 to 500; 6 = 500 or more.

Model Fits with Additional Endogenous Variables

To illustrate the improvement in explained variance in achievement occurring when 'endogenous' variables (those which are part of the correlated set of current student dispositions connected with achievement) are added to a model of combined mathematics, two additional variables were fitted as a further model: self-efficacy in mathematics and anxiety towards mathematics (see Appendix B, Section B.1). Their interrelationship with current academic achievement is such that they may be also considered as joint outcomes (in other words they are endogenous), both affecting, and affected by, the students' recent academic performance. The improvement in the explained variation is 4.9% at the cluster level and 14.7% at the individual level over that of the final model, giving resultant values of 83.8% and 44.3%, respectively. The marked improvement at the individual level illustrates the way in which the 'explained' variation can be apparently improved over the more stable student characteristics already fitted, but at the expense of the introduction of endogenous variables in a circular fashion. Because of these conceptual difficulties, the final model presented above involves a selection of explanatory variables which avoid the use of any joint outcomes as regressors.

MODELS OF READING LITERACY

Development of the Model

As with the models of combined mathematics, separate models of reading literacy were initially fitted. The parameter estimates, standard errors, and significance tests are shown in Table 5.9 (student-level variables) and Table 5.10 (school-level variables). The logarithm of the index of books in the home was used as before. Factors with more than two categories were tested by means of omnibus tests of the deviance change referred to the χ^2 distribution. All these models showed significant improvements over the null random intercept model.

Table 5.9. Achievement in Reading Literacy: All Student-Level Variables Tested as Separate Models by Addition to the Null Random Intercept Model

	Parameter	*SE*	*Test Statistic*	*df*	*p–value*
Gender: female–male	28.534	3.645	t = 7.828	158	<.001
Socioeconomic status	1.347	0.103	t =13.047	37	<.001
Lone parent: dual–lone	24.198	3.993	t = 6.060	161	<.001
Number of siblings			Ddiff = 63.473	4	<.001
none–one	–14.602	5.336			
two–one	–10.314	4.258			
three–one	–15.688	4.467			
four or more–one	–31.559	4.657			
Log books index	48.427	3.069	t = 15.779	75	<.001
Home educational resources			Ddiff = 96.984	2	<.001
low–high	–34.454	4.452			
medium–high	–17.059	3.214			
Absence			Ddiff = 109.651	2	<.001
1 or 2 days–no days	–15.222	3.009			
3 days or more–no days	–44.914	4.630			
Grade level			Ddiff = 268.033	3	<.001
Grade 8–Grade 9	–81.397	8.138			
Grade 10–Grade 9	42.600	3.889			
Grade 11–Grade 9	34.083	3.606			

Note: Ddiff = deviance difference, tested using the χ^2 distribution. Categorical variables with two or more levels are in italics.

Table 5.10. Achievement in Reading Literacy: All School-Level Variables Tested as Separate Models by Addition to the Null Random Intercept Model

	Parameter	*SE*	*Test Statistic*	*df*	*p–value*
Size (stratum)			Ddiff = 13.892	2	<.001
small–large	–44.256	16.735			
medium–large	–21.679	7.428			
Sector			Ddiff = 48.289	2	<.001
comm/comp–secondary	–26.908	7.622			
vocational–secondary	–51.522	7.128			
Disciplinary climate	25.88	8.498	t = 3.046	143	.003
Percent JCE fee waiver	–1.867	0.160	t = –11.694	143	<.001

Note: Ddiff = deviance difference, tested using the χ^2 distribution. JCE = Junior Certificate Examination. Categorical variables with two or more levels are in italics.

All the variables were then entered simultaneously as main effects and non-significant variables removed in sequence, with the least significant term removed each time and the retained terms re-evaluated. As in the case of the models for combined mathematics, school size (stratum) and school sector were each removed in turn. None of the omnibus tests of the student-, cross- and school-level two-way interactions were significant, suggesting that a main effects model is adequate. A linear model for the variates apart from log books index was obtained, since none of the tests of the additional polynomial terms was statistically significant. Finally, there were no significant random coefficients (tested by deviance differences when added to the model) resulting in the selected model having a random intercept. This indicates that the effects of the student-level variables do not vary appreciably across schools.

The Final Model

The parameter estimates, standard errors and significance tests for the final model of reading literacy are shown in Table 5.11. In contrast to combined mathematics, significance tests are reported for all main effects, since no interactions are present in the model which would necessitate the inclusion of the corresponding marginal main effects. Compared to the parameter estimates for the variables fitted as separate models (given in Tables 5.9 and 5.10), the changes are similar to those found for combined mathematics; e.g., students in lone-parent families and the estimates associated with values of number of siblings other than one show stronger deficits in scores compared to their respective reference categories. School size (stratum) and sector were both eliminated from the model, as in the case of mathematics, indicating that their originally explained variation (when fitted separately) is now covered by the remaining terms of the final model.

Table 5.11. Final Model of Achievement in Reading Literacy

	Parameter	*SE*	*Test Statistic*	*df*	*p–value*
Intercept	471.838	6.295			
Student–Level Main Effects					
Gender: female–male	21.466	3.051	t = 7.037	75	<.001
Socioeconomic status	0.798	0.096	t = 8.303	44	<.001
Lone parent: dual–lone	13.255	3.708	t = 3.574	74	.001
Number of siblings			Ddiff = 30.430	4	<.001
none–one	–7.693	4.707			
two–one	–8.413	3.878			
three–one	–11.2	4.023			
four or more–one	–19.302	4.279			
Log books index	32.137	2.952	t = 10.887	95	<.001
Home educational resources			Ddiff = 21.451	2	<.001
low–high	–14.505	4.116			
medium–high	–7.673	2.865			
Absence			Ddiff = 82.393	2	<.001
1 or 2 days–no days	–10.04	2.743			
3 days or more–1 or 2 days	–35.34	4.087			
Grade level			Ddiff = 328.496	3	<.001
Grade 8–Grade 9	–66.088	7.465			
Grade 10–Grade 9	33.536	3.552			
Grade 11–Grade 9	37.253	3.355			
School–Level Main Effects					
Disciplinary climate	18.397	4.633	t = 3.971	142	<.001
Percent JCE fee waiver	–1.056	0.123	t = –8.592	142	<.001
Variance Components					
Intercept	219.532				
Residual	4330.709				

Note: Ddiff = deviance difference, tested using the χ^2 distribution. JCE = Junior Certificate Examination. Variables dropped during model-building (in sequence): School size (stratum) and school sector. Categorical variables with two or more levels are in italics.

The null random intercept model has an ICC that gives a percentage of the total variation in achievement attributable to the cluster level (i.e., schools), of 19.8%. Fitting the submodel that omits the school variables gives explained variances of 61.3% at the cluster-level and 31.2% at the individual-level (classes and students). The addition of the school variables disciplinary climate in mathematics classes and Junior Certificate Examination fee waiver to obtain the final model results in increases in explained variance of 20.1% and 4.2% at the cluster and individual levels, respectively. The final model, therefore, explains 81.4% at the cluster-level and 35.4% at the individual level (see also Figure 5.1).

Contributions of the Explanatory Variables to Fitted Scores in Reading Literacy

The intercept has a value of 471.8 score points and reflects the choice of reference categories for the factors, as well as the mean centred variables, socioeconomic status (SES), disciplinary climate, and Junior Certificate Examination fee waiver. As with mathematics, the intercept is the fitted score for a (hypothetical) male student, of average SES, in a lone-parent family, with one sibling, with zero to ten books in the home, high home educational resources, 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 average fee waiver level for the Junior Certificate Examination. The insertion of the contributions of other combinations to the fitted score (calculated using their parameter estimates) produces fitted scores for different values of the student and school variables.

Examples of score contributions from the categorical variables can be read directly from the table of parameter estimates. The significant gender difference is 21.5 points (close to one-quarter of a standard deviation) in favour of females. Students in the medium category of home educational resources scored 7.7 points (under one-tenth of a standard deviation) lower than those in the high category, while students in the low home educational resources category students scored 14.5 points (one-sixth of a standard deviation) lower than their high category counterparts.

All the continuous variables appear linearly as main effects and so the parameter estimates give the change in student scores corresponding to a one unit change in the corresponding variate. As in the mathematics model, examples are provided for students at the mean of the lower, middle, and upper thirds of the population on a given variate.

The contributions to fitted scores for example values of SES are given in Table 5.12. The top row gives the means of the three groups of SES values (low, middle, and upper third). The second row consists of these values centred by subtraction of the overall weighted mean of 48.3 SES units. The final row provides the contributions to the fitted scores suggested by the model. Students at the mean of the medium and low groups of the SES variable are expected to score 14.9 (one-sixth of a standard deviation) and 28.8 points (one-third of a standard deviation) less than a student at the mean of the high group, respectively.

Table 5.12. Contribution to Fitted Scores in Reading Literacy for Example Values of Student Socioeconomic Status (SES)

	SES		
	Low	*Medium*	*High*
Raw SES Score	30.54	47.88	66.59
Centred SES Score	–17.95	–0.61	18.10
Contribution to Student Score	–14.33	–0.48	14.45

Note. SES values are centred by subtracting the weighted SES mean (48.34) before the fitted values are calculated.

No interaction of log books index was required in the reading model so Table 5.13 presents the score contributions for each level of the books index alone. As in the case of combined mathematics, there is a suggestion of a ceiling to the improvements in student scores corresponding to higher amounts of books in the home. Compared to a student coming from a home with zero to 10 books, students show average improvements of 22.3, 35.3, 44.6, 51.7, and 57.6 score points as the books index increases.

Table 5.13. Contribution to Fitted Scores in Reading Literacy for Books in the Home

	Book Index					
	None to 10 (1)	*11 to 25 (2)*	*26 to 100 (3)*	*101 to 200 (4)*	*201 to 500 (5)*	*500+ (6)*
Contribution to Student Score	0.00	22.28	35.31	44.55	51.72	57.58

Table 5.14 displays the contribution to fitted scores for three examples of the school variable disciplinary climate in mathematics classes. Students who attend a school at the mean of the medium disciplinary climate group score 24.2 points (just over a quarter of a standard deviation) lower than students at the mean of the high level, while the deficit for a student at the mean of the low group is 49.9 points (over half a standard deviation) compared to a student in the high group.

Table 5.14. Contribution to Fitted Scores in Reading Literacy for Example Values of Disciplinary Climate (in Mathematics Classes)

	Disciplinary Climate		
	Low	*Medium*	*High*
Raw Disciplinary Climate Score	–1.09	0.29	1.62
Centred Disciplinary Climate Score	–1.35	0.03	1.36
Contribution to Student Score	–24.91	0.54	24.95

Note. Disciplinary climate values are centred by subtracting the weighted disciplinary climate mean (0.2651) before the fitted values are calculated.

Score contributions for example values of Junior Certificate Examination fee waiver are given in Table 5.15. It can be observed that, in comparison to a student in a school at the mean of the high grouping, a student in a school at the mean of the medium level scores 21.9 points (around a quarter of a standard deviation) more, on average. Students in a school at the mean of the low group score 35.0 points (two-fifths of a standard deviation) higher than students in a school at the mean of the high group.

Table 5.15. Contribution to Fitted Scores in Reading Literacy for Example Values of Fee Waiver for the Junior Certificate Examination

	JCE Fee Waiver		
	Low	*Medium*	*High*
Raw JCE Fee Waiver Score	10.67	23.07	43.80
Centred JCE Fee Waiver Score	–15.17	–2.77	17.96
Contribution to Student Score	16.02	2.92	–18.97

Note. JCE fee waiver values are centred by subtracting the weighted JCE fee waiver mean (25.84) before the fitted values are calculated.

MODELS OF SCIENCE

Development of the Model

Following the pattern of model development for combined mathematics and reading literacy, separate models for each explanatory variable were first fitted for science. Tables 5.16 and 5.17 show the resulting parameter estimates, standard errors and significance tests for each model at the student and school levels. With the exception of gender, all the variables provided significant improvements over the null random intercept model. The variable study of science at Junior Certificate level is included as an additional explanatory variable in the context of the models of science.

The variables were all entered simultaneously as main effects terms. Gender was retained at this point to see if any differences between females and males might become significant, following conditioning on the other variables. The backward elimination of the least (non-)significant variables proceeded until a model was obtained in which all terms were required. School size (stratum) and sector were removed in sequence, following the pattern observed in the combined mathematics and reading literacy models. A main effects model was found to be adequate as the omnibus tests of student-, cross- and school-level two-way interactions were not statistically significant. As with the models for mathematics and reading, no random coefficients were significant for science and a random intercept model was sufficient, suggesting a constant effect of the student variable across the school clusters.

Table 5.16. Achievement in Science: All Student-Level Variables Tested as Separate Models by Addition to the Null Random Intercept Model

	Parameter	*SE*	*Test Statistic*	*df*	*p–value*
Gender: female–male	–2.884	3.911	t =–0.737	422	.461
Socioeconomic status	1.649	0.119	t =13.825	28	<.001
Lone parent: dual–lone	24.296	4.267	5.694	1141	<.001
Number of siblings			Ddiff = 57.040	4	<.001
none–one	–17.059	5.864			
two–one	–11.414	4.638			
three–one	–17.481	4.765			
four or more–one	–33.397	5.080			
Log books index	59.934	3.156	t = 18.993	423	<.001
Home educational resources			Ddiff = 117.692	2	<.001
low–high	–41.366	4.518			
medium–high	–23.77	3.709			
Absence			Ddiff = 103.802	2	<.001
1 or 2 days–no days	–19.203	3.452			
3 days or more–no days	–47.725	6.151			
Grade level			Ddiff = 226.471	3	<.001
Grade 8–Grade 9	–76.216	10.575			
Grade 10–Grade 9	42.620	4.781			
Grade 11–Grade 9	31.008	4.085			
Science for JCE: no–yes	–61.414	5.659	t = –10.853	195	<.001

Note: Ddiff = deviance difference, tested using the χ^2 distribution. Categorical variables with two or more levels are in italics.

Table 5.17. Achievement in Science: All School-Level Variables Tested as Separate Models by Addition to the Null Random Intercept Model

	Parameter	*SE*	*Test Statistic*	*df*	*p–value*
Size (stratum)			Ddiff = 13.085	2	<.001
small–large	–41.825	16.080			
medium–large	–19.026	6.960			
Sector			Ddiff = 38.159	2	<.001
comm/comp–secondary	–20.691	7.536			
vocational–secondary	–43.718	7.086			
Disciplinary climate	16.857	8.195	t = 2.057	143	.041
Percent JCE fee waiver	–1.761	0.149	t = –11.846	143	<.001

Note: Ddiff = deviance difference, tested using the χ^2 distribution. JCE = Junior Certificate Examination. Categorical variables with two or more levels are in italics.

The Final Model

The final model of science is shown in Table 5.18. As in the reading literacy model, the lack of any interactions allows significance tests of all the main effects included in the final model. A comparison with the parameter estimates for the separately fitted models for each variable (see Tables 5.16 and 5.17) shows that the deficit in scores for females compared to males has increased and that it is statistically significant in the presence of the other model terms. Other variables retained in the final model vary in whether they show a decrease or increase of their estimated parameters compared to when fitted as separate models. Again, the selected variables appear to explain sufficient variation in scores encompassed by school size (stratum) and sector, such that the latter two variables could be eliminated.

The proportion of total variation in scores suggested to occur at the school level is 13.3%, as given by the ICC of the null random intercept model. Omitting the school variables from the final model to give a model with the student-level variables explains 66.6% of cluster variation and 29.0% of individual-level (class and student) variation. Adding in the school variables produces improvements of 13.6% and 2.2% of the explained variances at the cluster and individual level, respectively. This results in the final model which explains 80.2% of the between-cluster variation and 31.2% of within cluster variation (see also Figure 5.1).

Contributions of the Explanatory Variables to Fitted Scores in Science

A value of 477.3 score points is obtained for the intercept of the final model. This is the fitted score with the factors having the reference categories indicated and using the mean centred values for SES, disciplinary climate and Junior Certificate Examination fee waiver. It represents a hypothetical student similar to that represented by the intercept in the mathematics and reading models with the addition of the study of science variable, i.e., male, average SES, from a lone-parent family with one sibling, none to 10 books at home, high home educational resources, no absence in the two weeks preceding the PISA assessment, in Grade 9, who studies/studied science for the Junior Certificate, is at a school with average disciplinary climate and average fee waiver. The parameter estimates enable fitted scores to be calculated for the other types of student in the manner described above.

Table 5.18. Final Model of Achievement in Science

	Parameter	*SE*	*Test Statistic*	*df*	*p–value*
Intercept	477.276	6.380			
Student–Level Main Effects					
Gender: female–male	–7.583	3.091	t = –2.453	404	.015
Socioeconomic status	0.921	0.115	t = 8.034	28	<.001
Lone parent: dual–lone	9.741	3.814	t = 2.554	812	.011
Number of siblings			Ddiff = 24.238	4	<.001
none–one	–10.849	5.171			
two–one	–9.3	4.132			
three–one	–12.646	4.230			
four or more–one	–19.567	4.477			
Log books index	39.731	3.116	t = 12.751	589	<.001
Home educational resources			Ddiff = 28.328	2	<.001
low–high	–16.999	4.246			
medium–high	–12.468	3.387			
Absence			Ddiff = 62.273	2	<.001
1 or 2 days–no days	–10.567	3.093			
3 days or more–1 or 2 days	–34.358	5.581			
Grade level			Ddiff = 226.066	3	<.001
Grade 8–Grade 9	–61.808	9.813			
Grade 10–Grade 9	32.281	4.300			
Grade 11–Grade 9	35.939	3.686			
Study of Science: no–yes	–38.153	5.043	t = –7.565	115	<.001
School–Level Main Effects					
Disciplinary climate	15.505	4.840	t = 3.204	142	<.001
Percent JCE fee waiver	–0.844	0.121	t = –7.000	142	<.001
Variance Components					
Intercept	167.401				
Residual	5493.798				

Note: Ddiff = deviance difference, tested using the χ^2 distribution. JCE = Junior Certificate Examination. Variables dropped during model-building (in sequence): School size (stratum) and school sector. Categorical variables with two or more levels are in italics.

The parameter estimates for categorical variables provide a direct interpretation of differences for each indicated category relative to the reference category for a factor. The now significant gender difference (compared to the separate model for gender alone, which was not significant) suggests that females tend to score 7.6 points lower than males (just under one-tenth of a standard deviation), adjusting for the other variables in the model. Students who do not study science show a deficit of 38.2 score points (over two-fifths of a standard deviation) compared to students who take the subject at school.

Students with none or two siblings score 10.8 and 9.3 points (just under one-eighth of a standard deviation) lower than students with one sibling, respectively. The deficit is larger for students with three and four or more siblings: 12.6 points (just over one-eighth of a standard deviation) and 19.6 points (just under a quarter) lower, respectively, than those with one sibling.

The continuous variables, as with the final reading literacy model, all appear linearly as main effects. Therefore, the parameter estimates reflect the change in the fitted score per unit of each variate. Again, examples are provided for students who report a covariate value at the mean of the lower, middle, and upper thirds of the population.

Examples of the fitted score contributions for values of SES are given in Table 5.19 with the top row indicating the SES values of means of the three groups. The middle row is centred by subtraction of the overall SES mean, and the final row gives the contributions to fitted scores. A student who has a reported SES value at the mean of the medium and low groups scores on average 17.2 and 33.2 points (one-fifth and two-fifths of a standard deviation) lower, respectively than their counterparts at the mean of the high SES group.

Table 5.19. Contribution to Fitted Scores in Science for Example Values of Student Socioeconomic Status (SES)

	SES		
	Low	*Medium*	*High*
Raw SES Score	30.54	47.88	66.59
Centred SES Score	–17.95	–0.61	18.10
Contribution to Student Score	–16.54	–0.56	16.67

Note. SES values are centred by subtracting the weighted SES mean (48.34) before the fitted values are calculated.

Table 5.20 shows the contribution to fitted scores for the index of books in the home in the final science model which, as with reading, has only a main effect and can be presented by itself. The asymptote for the increase in scores as the index of books in the home increases is found once more, as occurred with the mathematics and reading models. Students with a value corresponding to each point on the books in the home scale show the following pattern of increasing improvement in scores over students with no books: 27.5, 43.7, 55.1, 63.9 and 71.2 score points. These correspond to differences ranging from one-third to four-fifths of a standard deviation. Similar to the fitted values for reading, a ceiling effect is suggested.

Table 5.20. Contribution to Fitted Scores in Science for Books in the Home

	Book Index					
	None to 10 (1)	*11 to 25 (2)*	*26 to 100 (3)*	*101 to 200 (4)*	*201 to 500 (5)*	*500+ (6)*
Contribution to Student Score	0.00	27.54	43.65	55.08	63.94	71.19

Score contributions for the school variable disciplinary climate are shown in Table 5.21. These suggest that, compared to a student in a school at the mean of the high group, those attending schools at the means of the medium and low groups tend to score 20.6 and 42.3 points less, respectively. These correspond to one quarter and one-half of a standard deviation on the science scale.

Table 5.21. Contribution to Fitted Scores in Science for Example Values of Disciplinary Climate (in Mathematics Classes)

	Disciplinary Climate		
	Low	*Medium*	*High*
Raw Disciplinary Climate Score	–1.09	0.29	1.62
Centred Disciplinary Climate Score	–1.35	0.03	1.36
Contribution to Student Score	–21.00	0.45	21.03

Note. Disciplinary climate values are centred by subtracting the weighted disciplinary climate mean (0.2651) before the fitted values are calculated.

Table 5.22 shows the contribution to fitted scores for examples of the Junior Certificate Examination fee waiver variable. Students in schools at the mean of the medium and low groups of fee waiver are expected to score 17.5 and 28.0 points (one-sixth and one-third of a standard deviation) higher, respectively, than students in schools at the mean of the high fee waiver grouping.

Table 5.22. Contribution to Fitted Scores in Science for Example Values of Fee Waiver for the Junior Certificate Examination

	JCE Fee Waiver		
	Low	*Medium*	*High*
Raw JCE Fee Waiver Score	10.67	23.07	43.80
Centred JCE Fee Waiver Score	–15.17	–2.77	17.96
Contribution to Student Score	12.80	2.34	–15.16

Note. JCE fee waiver values are centred by subtracting the weighted JCE fee waiver mean (25.84) before the fitted values are calculated.

CONCLUSION

The gender differences observed in the descriptive tables of Chapter 4 (e.g., Table 4.1) were confirmed in both the separate models and in the final model for combined mathematics that conditions on a range of student- and school-level explanatory variables. No gender interactions were suggested by the model-building process. The final model suggests a deficit of 24.2 points (three-tenths of a standard deviation) for female students compared to males, adjusting for the other variables in the model. The interaction between the logarithm of books in the home index and absence from school indicates the strength of the relationship between books at home and achievement. At least 42.4 points (half of a standard deviation) separates students with no books from those with 500 or more. This difference is 69.4 points (four-fifths of a standard deviation) for students who reported zero absences in the two weeks prior to taking the PISA test. The effect of three or more days absence is at least 13.0 score points (close to one-sixth of a standard deviation), depending on the number of books in the home. The difference between students from single and dual-parent families is 15.7 points (one-sixth of a standard deviation) in favour of the latter group.

Two school-level variables were retained by the model selection process: the average disciplinary climate of a school, as reported by students, and the percentage of students eligible for the Junior Certificate Examination fee waiver. The difference between students in schools scoring at the mean of the middle and lower groups of schools on the disciplinary climate variable, compared to the higher group was a deficit of 25.4 points and 52.1 points (one-third and three-fifths of a standard deviation), respectively. The score differences estimated for the medium and low levels of percent fee waiver, compared to the high level, were 16.6 points and 26.5 points (one-sixth and one-third of a standard deviation), respectively. The addition of these two variables to the student-level variables in the final model raises its explained variance by 16.5% at the cluster-level and 2.7% at the class/student level.

The addition of mathematics self-efficacy and anxiety about mathematics improved the fit further, compared to the final model chosen (an additional 4.9% variance explained at the cluster (school) level and 14.7% at the individual (class and student) level), but raised conceptual difficulties about their status as explanatory variables, since they might be better considered as outcomes of the education process.

The final model for reading literacy indicates a significant gender difference; females scored 21.5 points (one-quarter of a standard deviation) higher than males, following adjustment for the other variables in the model. The gain in achievement associated with increasing numbers of books in the home followed a similar pattern to that observed in the model of combined mathematics, though a main effect for that variable (no interaction with absence or any other variable) was sufficient in the case of reading. Students in lone-parent families had a deficit of 13.3 points (close to one-sixth of a standard deviation) compared to students in dual-parent families, which is similar to the deficit found for combined mathematics. The score contributions for example values of student-level SES suggests a difference of 28.8 points (one-third of a standard deviation) between students at the mean of the low SES group and the high SES group.

Disciplinary climate and percent fee waiver were the two school-level variables selected by the backward elimination procedure (and are in the final reading model). The fitted contributions for the example values suggest their associations with achievement are similar to those found for combined mathematics. Their addition to the student-level variables for the final model in reading increases its explained variance by 20.1% at the cluster level (i.e., between schools) and 4.2% at the individual level corresponding to classes and students.

As in the case of reading literacy, the final model for science requires no interactions; main effects are sufficient. Once the other variables in the model are included, the gender difference increases to a statistically significant level with females scoring 7.6 points (just under a tenth of a standard deviation) lower than males on average. This is of a smaller magnitude, though in the same direction, as in the final model for combined mathematics. The tendency for scores to increase as students report more books in the home is similar to that found for reading literacy, but with a higher asymptote. The difference between scores of students with zero to 10 books and those with 500 or more books is 71.2 points (around four-fifths of a standard deviation) and is similar to the maximum value found in the mathematics model where it interacted with frequency of absence from school. The difference between students in lone-parent and dual-parent families is 9.7 points (one-tenth of a standard deviation) in favour of students in dual-parent families and somewhat smaller than that found in the other domains. The size of the SES association with science achievement falls between that found in the mathematics and reading models (the SES association is smallest in the case of reading). Students who are at the mean of the low SES group score 38.2 points (two-fifths of a standard deviation) lower in science than students at the mean of the high SES group. Students who studied science at Junior Certificate level scored 38.2 points (two-fifths of a standard deviation) higher than students who did not.

The same two school-level variables as the other two domains, disciplinary climate and percent fee waiver, are found in the final model of science. Their impact is in the same direction, though the absolute size of the associations is somewhat smaller than that found in mathematics and reading. Adding both the student-level variables of the final model provides gains of 13.6% and 2.2% at the cluster (school) and individual (class and student) level, respectively.

The comparatively low intra-cluster correlations (ICCs) for all three domains indicate, as found in PISA 2000, that the Irish school system is comparatively homogenous at the level of the school. The sample design of PISA, which entails a random selection of 15-year-olds across different classrooms, does not allow the partitioning of within-school achievement variance into between-classroom and within-classroom components. This results in a single large within-school variance component that reflects both class- and student-level variation.

The between-school variance components for Irish students' performance on combined mathematics, reading and science are 16.7%, 22.5%, and 16.2%, respectively[45], The values for combined mathematics and reading literacy but not for science were somewhat lower in PISA 2000.[46] Values for both 2000 and 2003 are also similar to the values reported by Smyth (1999), whose estimates are derived from a composite value based on six Junior Certificate Examination subjects. The PISA 2000 and PISA 2003 values are lower than for studies using grade-based samples such as the Third International Mathematics and Science Study (TIMSS 1995), which reported ICCs of 50% for mathematics and 38% for science (Martin et al., 2000). In the IEA Reading Literacy Study (IEA/RLS), the corresponding value for reading was 48% (Postlethwaite, 1995). These differences are likely to reflect differences arising from a random within-school sample design as in PISA (giving school and class + student variance components) and a design involving the sampling of intact classes as in TIMSS mathematics (giving school + class and student components).

Multilevel models were presented for each domain. The models explain at least 78.8% of between-school variance and 29.6% of within-school variance in all three domains. As in PISA 2000, the largest explained variance values were found for reading literacy. The addition of two variables which are highly correlated with current achievement to the final combined mathematics model (anxiety towards mathematics and mathematics self-efficacy,) improved its explained variation by 16.5% and 2.7% between and within schools, respectively. This additional 'explained' variation is at the expense of considerable conceptual difficulty since the variables are measured concurrently with achievement. In fact, the students completed the PISA test immediately prior to the questions used to construct the efficacy and anxiety scales, leading to further possible inflation of their associations. Such gains for adding endogenous variables are similar to those reported when a similar exercise was carried out for a model of PISA 2000 reading literacy in which early school leaving intent and attitude to reading were added as regressors (Sofroniou, May 2004). Furthermore, self-efficacy and anxiety can be viewed as joint outcomes with achievement that may be better dealt with by more complex multivariate multilevel models which allow more than a single response variable (e.g., Hox 2002). Because of these interpretive and theoretical problems, the final models reported here avoid such highly interrelated variables in the multilevel regression procedures used. This lower emphasis of the explanatory utility of variables based on self-reports differs from that of the OECD PISA 2003 initial report (OECD, 2004b). It seems likely that the conceptual circularity entailed by their inclusion as regressors would serve to negate any useful improvement in the fit of the model, indicated here by the increase in explained variance at each level.

The substantial proportion of between-school variance that is explained by the student-level variables when they are fitted without school-level variables (e.g., 62.3% for combined mathematics) indicates that considerable variation in school performance is due to differences in student composition (see Hox, 2002). Measures for school-level variables are provided by the addition of the disciplinary climate and Junior Certificate Examination fee waiver variables (explaining a further 16.5% of between-school variation in mathematics). The latter illustrates the contextual effect of the school-level SES-related variable over and above the individual student measures used in the model.

A wider range of interactions was examined in PISA 2003 compared to PISA 2000: student-, cross- and school-level two-way interactions were all examined in 2003, rather than just gender interactions, as was the case in 2000. The final models, reported here, for all

[45] The ICCs as estimated from the PISA 2003 models of achievement for Ireland (using full maximum likelihood) are close to these values, at 14.6%, 19.8% and 13.3%, respectively.
[46] The corresponding values reported for PISA 2000 are 11.4%, 17.8% and 14.1%.

three PISA domains, are very similar to one another, with the exceptions of the addition of the log books index interaction term with absence to the model of mathematics, and a domain specific variable, study of Junior Certificate science, to the science model. In that sense, they are more similar to one another than those reported for PISA 2000 (Shiel et al., 2001). The model used to scale students' achievement data in PISA 2003 generated scale scores for all students in the minor domains, even though six out of every 13 had not attempted items in any given minor domain. In PISA 2000, scale scores were not computed for those students who had not attempted items from the minor domains, resulting in scores in mathematics and science for just five in nine students. Therefore, the sample sizes are now the same for each domain, with the result that the complexities of the final models in each domain are similar. It may be noted that students who did not attempt items on the assessments of reading literacy and of science in 2003 had mean scale scores in both of these domains that were significantly lower than those who had attempted items, by just over one-eighth of a standard deviation in both cases (12.2 score points in the case of reading and 11.6 points in the case of science). It may be the case that this procedure has also produced patterns of associations amongst the explanatory variables and the achievement scores that are artificially similar across domains, reflected in the near identical models for each academic area reported in the present chapter. This is an area that would benefit from further research.

Significant gender differences are present in all three final models, in favour of males for combined mathematics and science, and in favour of females for reading literacy. These are of a similar magnitude for combined mathematics and reading, while the value for science is considerably smaller (having changed from non-significant when originally fitted separately). The precise combination of variables corresponding to the significant gender difference in science may be worth pursuing in future research.

The importance of books in the home was once again confirmed; its contribution to achievement is stronger for mathematics in the case of students who reported no recent absence from school. Consistent deficits were shown for students coming from lone-parent backgrounds, compared to those from dual-parent households, though the borderline significant interaction between lone-parent status and gender in PISA 2000 mathematics was not replicated.

The inclusion of information about whether or not each student studied science for the Junior Certificate showed a significant difference in science achievement in favour of those who did. However, investigation of this variable with achievement in mathematics and reading suggests that it may be acting as a proxy for low achievement, since similar achievement differences are also found for these two domains (see Table 4.22). Identifying reasons why students in Ireland do not study science at Junior Certificate level is worthy of investigation. As a substantial minority (9.9%) of Irish students who participated in PISA did not study science, the question may be raised as to whether at least some of these students would benefit in the future by having encountered at least some science in their post-primary school curriculum.

The impact of SES at the individual level found in PISA 2000 is again confirmed in 2003, with the addition of a useful school-level deprivation context measure that is easily collected and policy relevant, namely percent of students entitled to the Junior Certificate Examination fee waiver. This appears to be a more informative measure than the Department of Education and Science dichotomous school disadvantage status used in the models for PISA 2000 (Shiel et al., 2001), as it captures the more continuous relationship of SES density of peers in the school context with achievement. It is also notable that, in contrast to PISA 2000, school sector does not appear in any of the final models for 2003. This suggests that the current model better explains the achievement differences across secondary, vocational, and community/comprehensive schools since the variation corresponding to school type is accounted for.

As noted in Chapter 4, the disciplinary climate in variable, which appears in all three models, may be associated with school management practices. The measure is an aggregate of student self-reports within a given school, and students have been sampled at random across different classes and grade levels. The development of a more general measure of school disciplinary climate should be considered.

In addition to SES, the impact of a range of student variables observed in 2000 is replicated, including the number of books in the home, family size and frequency of absence from school. Substantial differences between grade levels were found and including this variable improved the fit of the model considerably. Differences associated with the grade level variable, however are difficult to interpret. Students can be above or below the modal grade level (Third Year) for a variety of reasons. Further, the Transition Year Programme is not available to, or taken by, all students. Of relevance here is that, in Chapter 4, a comparison of students at different grade levels revealed achievement differences in the region of one-third of a standard deviation between students in Fifth year and students in Fourth (Transition) year (Table 4.19). However, the models reported in this chapter indicate that achievement differences between Fourth and Fifth years, after adjusting for the other variables, are small (just two to three scale points).

6

Curriculum and Assessment in Ireland and Performance on PISA 2003

This chapter examines (i) relationships between performance in mathematics, English and science on the Junior Certificate Examination and performance on PISA 2003, (ii) content overlap between PISA mathematics and Junior Certificate mathematics, (iii) relationships between the expected familiarity of students with items in the PISA 2003 mathematics assessment and their performance on PISA 2003 mathematics, and (iv) the distribution of students at each Junior Certificate syllabus level in mathematics and English across PISA combined mathematics and reading literacy proficiency scales, respectively.

Before describing the outcomes of these analyses, the Junior Certificate syllabus in mathematics current at the time of the PISA 2003 assessment (referred to here as the 'current mathematics syllabus') is described in brief. The performance of students in mathematics, English, and science in the 2003 Junior Certificate Examination is also described.

In the recent report of the Task Force on the Physical Sciences (2002), concerns about students' mathematical competencies were linked to a declining rate of uptake of the physical sciences. The Report notes that 17% of students in the 2001 Leaving Certificate mathematics examination obtained grade E or lower across the three syllabus levels and adds:

> Students' perception of the difficulty of mathematics and their poor performance in the subject both act as barriers to participation and success in the sciences at second and third level[s]. The risk in not addressing the problem with mathematics is that of undermining reform in science education. (p. 12)

Indeed, one of the recommendations of the Task Force is that the Department of Education and Science should undertake a review of mathematics that not only looks at academic factors, but also at social and behavioural factors.

Given that mathematics is the major focus of PISA 2003, and the current Junior Certificate mathematics syllabus was examined for the first time in 2003, it is also a valuable opportunity to place Junior Certificate mathematics performance in an international context. PISA 2006 will offer an opportunity to examine Junior Certificate science in an international context, since 2006 will see the first examination of the revised Junior Certificate science syllabus (introduced in September 2003).

THE JUNIOR CERTIFICATE MATHEMATICS SYLLABUS

Principles of the Junior Certificate Programme

When considering the outcomes reported in this chapter, it is useful to bear in mind that the Junior Certificate programme as a whole is based on three principles: (i) *breadth and balance* (the reinforcement and development of students' existing numeracy, literacy and oracy skills, emphasising social and environmental education, science and technology and modern languages); (ii) *relevance* (the curriculum should address both the immediate and prospective needs of young people, in the context of their cultural, economic and social environments); and (iii) *quality* (every young person should be challenged to achieve high standards with due regard to individual abilities and aptitudes, as well as to international comparisons) (Department of Education and Science/National Council for Curriculum and Assessment, 2000, p. 2).

Description of the Current Mathematics Syllabus

In 1994, the NCCA Mathematics (Junior Cycle) Course Committee was asked to review the Junior Certificate mathematics syllabus, with a view to introducing some amendments (Department of Education and Science/National Council for Curriculum and Assessment, 2002). Amongst other things, the review was to take into account the work of the NCCA with respect to the primary school mathematics curriculum and the content of, and achievement on, the Junior Certificate Examinations since 1990. The Committee was requested to identify major issues of concern regarding syllabus design, implementation and assessment, address some issues around Foundation level mathematics, draft aims and objectives for each of the three syllabus levels, and prepare teacher guidelines. The review drew from numerous sources, ranging from Chief Examiners' reports and meetings with the Irish Mathematics Teachers' Association, to a study of outcomes on international assessments. The limited remit of the review obviated any substantial changes to the content and style of the syllabus.

The current mathematics syllabus (Department of Education and Science/National Council for Curriculum and Assessment, 2000, 2002), which followed the syllabus review, has two aims which are common to all three syllabus levels:

1. To contribute to the personal development of students;
2. To help to provide them with the mathematical knowledge, skills and understanding needed for continuing their education, and eventually for life and work.

While the first aim is rather broad, the second may be interpreted to be consistent, to some extent, with the PISA 2003 definition of mathematics, which, as described in Chapter 1, emphasises mathematical understanding and its application in real-world contexts, espousing the type of mathematical skills that students will require and be able to build on as they progress to future adult participation in society.

The aims of the current Junior Certificate syllabus are further broken down into a series of objectives, which may be summarised as follows:

A. Recall of mathematical facts;
B. Competencies needed for mathematics activities (i.e., instrumental understanding);
C. Awareness of mathematics as a system that makes sense (i.e., relational understanding);
D. Application of mathematical knowledge;
E. Analysis of information, including that presented in unfamiliar contexts;
F. Ability to create mathematics for oneself (e.g., make informed guesses, and debate or discuss these);
G. Development of psychomotor skills to attain objectives (e.g., orderly presentation, appropriate use of calculators, constructions and diagrams);
H. Ability to communicate mathematics (e.g., describe their own working and reasoning in written or spoken form);
I. Appreciation of mathematics; and
J. Awareness of the history of mathematics.

Of the ten objectives, six (A, B, C, D, G and H) are 'assessment objectives', examined through questions on the Junior Certificate mathematics examination papers, while four (E, F, I, J) are not. It should be noted that mathematics education in Ireland, assessment by teachers of their own students' work does not feature as part of the formal assessment process. Again, the limited remit of the syllabus review obviated the introduction of novel assessment techniques, including coursework, which in turn has implications for the extent to which the content of the objectives could be modified or added to.

The specific objectives of the Junior Certificate mathematics syllabus do not appear to fit readily into the broader aims. The objectives, however, may be compared to the PISA 2003 framework in a general way. This comparison portrays the Junior Certificate mathematics syllabuses as being somewhat similar to the PISA approach to mathematics. Objectives A and B (recall of mathematical facts; instrumental understanding) are consistent with the assumption underlying the PISA framework that, by age 15, students have mastered very basic skills. Objective C (development of relational understanding) is consistent with the PISA view of mathematics as an integrated subject and that in real life, students need the conceptual understanding of procedures to know which to apply to solve a particular problem. Objective H (communication of mathematics) applies to many of the PISA mathematics items where students are asked to explain or justify their mathematical reasoning.

There are also notable differences. For example, the syllabus objectives are phrased in a more abstract way than the knowledge and skills described in the PISA framework. In this regard, PISA offers a clear way of operationalising or mapping assessment objectives onto test items. Further, while PISA describes the domain of mathematics in terms of concepts, processes, and situations (i.e., it is a multidimensional framework), the Junior Certificate syllabus focuses more on procedures and concepts (i.e., it has fewer dimensions). The four objectives which are not assessed are likely to receive less emphasis in instruction than those which are. The real-life approach to mathematical problem-solving in PISA implies that the ability to solve problems in novel contexts is an important prerequisite for many of the items. This skill is not immediately apparent in any of the assessment objectives, although it is mentioned in the Objective E (which is not assessed). Objectives F and I (ability to create mathematics, development of an appreciation of mathematics) are also consonant with PISA's emphasis on the importance of fostering an interest and appreciation in mathematics as a valuable educational outcome in itself, but these two objectives are not assessed either.

The rationale provided for each syllabus level indicates that students taking the Junior Certificate at those levels will differ in the extent to which they can apply mathematical concepts and demonstrate understanding in a variety of contexts (Department of Education and Science/National Council for Curriculum and Assessment, 2000). At Higher level, the syllabus is geared towards students who are of above average mathematical ability, but not all students at this level will use academic mathematics in the future. This implies that a balance must be struck between challenging the most able students and encouraging those who are developing at a slightly slower pace. Thus at Higher level, the development of abstraction and generalisation skills is emphasised alongside the introduction of proofs. Ordinary level is geared towards average ability students and offers mathematics that is both meaningful and accessible, providing for the gradual introduction of more abstract ideas. The emphasis at this level is on the development of mathematics as a body of knowledge and skills that make sense and that can be used in many different ways. The Foundation level course is intended for those who are unsuited by Ordinary level mathematics. The course objectives involve developing knowledge and skills in basic mathematics and an awareness of the usefulness of mathematics. The emphasis is on building confidence, both in the students themselves, and in their involvement with mathematics as a discipline.

The structure of the current syllabus is very similar to the previous version, i.e., students at all three syllabus levels learn concepts associated with sets, number systems, applied arithmetic and measure, algebra, statistics, geometry, and functions and graphs. Higher and Ordinary level students also study trigonometry. One major content area that is covered in PISA but not in the Junior Certificate is probability and chance (as measured by the PISA Uncertainty subscale), which is reserved for Senior Cycle.[47]

A number of changes in syllabus content/assessment may be noted (see Department of Education and Science/National Council for Curriculum and Assessment, 2002, pp. 3-7). First, there is now no choice on the examination papers, resulting in an increase in coverage of the course. Second, the appropriate use of calculators is recommended, and calculators are now permitted in the Junior Certificate mathematics examinations. Third, refinements to teaching sequence and reduction in emphasis on transformational elements (returning to a more traditional congruency-based approach) have been made to geometry topics, particularly at Higher level. Fourth, logical argument and rigorous proof are emphasised to a greater degree at Higher level than previously. Fifth, only a subset of theorems at Higher level is examined (but all are on the course). Sixth, transformational geometry still features, but is treated as a separate sub-topic. Finally, changes include the removal of logarithms (which were seen to be too abstract and somewhat outdated), the reduction in scope of many topics to shorten the overall length of the course at Higher level, some simplification of algebra and coordinate geometry at Ordinary level, less emphasis on fractions, more emphasis on computing decimals, and increased emphasis on statistics, data handling and algebra at Foundation level.

It is suggested in the Junior Certificate mathematics teacher guidelines (Department of Education and Science/National Council for Curriculum and Assessment, 2000, p. 17) that the changes in content are accompanied by some changes in emphasis: an increase in emphasis on relational understanding, in the communication of one's reasoning and results, and in the appreciation of mathematics.

The specific content of the current Junior Certificate mathematics syllabus is described in Table 6.1. While the table does not give an indication of the detail in which topics are covered at each level, it can be seen that, while there is not a substantial difference in topics covered as the Higher and Ordinary level courses, the Foundation level course focuses mainly, but not exclusively, on 'social mathematics' (i.e., the types of mathematical concepts and operations that one is likely to encounter in everyday life, such as those involving money, percentages, area, and volume). The main difference between Higher and Ordinary level courses is in terms of the depth of topic coverage, particularly in the areas of algebra and geometry.

Oldham (2002) has commented that the current mathematics syllabus for the Junior Certificate (Department of Education and Science/National Council for Curriculum and Assessment, 2000) represents "essentially a minor update to deal with areas of the course that were giving rise to difficulties, rather than root-and-branch review" (p. 43). She also emphasises, however, that the incareer development programme accompanying the revisions is an attempt to move away from mechanistic approaches towards teaching for understanding, a change that is consonant with the philosophy underlying PISA mathematics.

The implementation of the current Junior Certificate mathematics syllabus has not seen an increase in performance on PISA mathematics in 2003 (relative to 2000) (Chapter 3). However, one should bear in mind that the impact of curricular change on students' achievements, if any, is likely to be slow and influenced by the impact of incareer training on

[47] It is interesting to note that the 1999 primary school mathematics curriculum includes a strand on data and chance. This represents a discontinuity in this area between primary, lower and upper second level mathematics education. It should be noted that, in the context of the 1994 syllabus review it was outside the remit of the NCCA mathematics course committee to suggest additional topics for Junior Certificate mathematics.

instructional techniques, mathematics textbooks associated with the revised syllabus, and the types of examination questions to be answered. Further, by the time students reach the beginning of second level schooling, they already have several years of learning mathematics at primary level. The mathematics component of the revised primary school curriculum (Department of Education and Science/National Council for Curriculum and Assessment, 1999) was not introduced in schools until 2002. Therefore, students participating in PISA 2003 will not have experienced the revised primary mathematics curriculum. Appendix 6 (Section A6.1) gives a brief description of the main changes to the mathematics curriculum at primary level.

Content and S tyle of the Junior Certificate Mathematics Examination Compared to PISA Mathematics

The aims and objectives of the Junior Certificate mathematics curriculum taken as a whole and the PISA mathematics assessment are somewhat similar in many respects; however, they also diverge. Divergence can arises for a number of reasons – it can be traced to the fact that the assessment objectives take higher priority than objectives which are not assessed. Further, Junior Certificate mathematics emphasises vertical mathematisation; i.e., developing increasingly complex mathematics concepts and skills in abstract contexts. PISA, in contrast, emphasises horizontal mathematisation; i.e., the application of mathematical concepts and skills to organise and solve a problem located in a real-life situation, and the abstraction of concepts and skills from these contexts (see Treffers, 1987).

Questions on the Junior Certificate mathematics examination are usually presented in three sections: part (a) (facts and skills) is intended as an introductory section; questions here involve recall of facts and simple, routine application of procedures. Hence they can be likened to the easier PISA Reproduction items. Part (b) is more demanding than part (a) as it may involve the understanding of concepts, the application of more complex procedures, or more steps to the solution; but in general the contexts are still routine and hence may be likened to the more difficult PISA Reproduction items. Part (c) questions are usually more complex and at times involve the extension of a routine procedure to an unfamiliar context. Hence, some part (c) questions can be likened to PISA Connections items. In general, Junior Certificate mathematics tend not to tap processes associated with items in the PISA Reflection cluster. Such items usually require significant extension of concepts and procedures and higher-level sustained mathematical reasoning or modelling, usually in the form of horizontal mathematising.

A comparison of PISA mathematics and the Junior Certificate mathematics examination reveals some marked differences in the manner in which problems are contextualised. In the Junior Certificate Examination papers, questions are usually presented in a purely mathematical and abstract context, almost always without redundant information. In the PISA assessment, on the other hand, questions are often embedded in rich real-life contexts, accompanied by texts and diagrams. In PISA, students are often required to discriminate between necessary and redundant information, as well as to actually formulate the problem, in order to solve it.

Table 6.1. Outline of Topics Covered at Higher, Ordinary, and Foundation Level Mathematics for the Junior Certificate: Revised (2000) Syllabus

Topic	*Sub-topic*	*H*	*O*	*F*
Sets	Elements, membership, universe, subset, null set, equality of sets	✓	✓	✓
	Venn diagrams	✓	✓	✓
	Set operations: intersection, union, complement	✓	✓	✓
	Set operations: difference	✓	✓	
	Set operations extended to three sets	✓	✓	
	Commutative property	✓	✓	✓
	Associative property	✓	✓	
	Distributive property	✓		
	Correct usage of brackets	✓	✓	
Number Systems	The set N: order, place value, sets of multiples, lowest common multiple [highest common factor Higher and Ordinary only]	✓	✓	✓
	The set N: sets of divisors, pairs of factors, prime numbers, cardinal number, rules for indices	✓	✓	
	The set N: addition, subtraction, multiplication and division, approximation	✓	✓	✓
	The set Z: positional order on the number line, addition	✓	✓	✓
	The set Z: order, addition, subtraction, multiplication, division, approximation	✓	✓	
	The set Q: decimals, fractions, percentages; addition, subtraction, multiplication and division	✓	✓	
	The set Q: square roots and reciprocals	✓	✓	
	The set Q: rounding, approximation	✓	✓	✓
	The set Q: ratio and proportion	✓	✓	
	The set Q: rules for indices e.g. $a^pa^q = a^{p+q}$, $a^p/a^q = a^{p-q}$, $a^0=1$	✓	✓	
	The set Q^+: common fractions (e.g., denominators 2, 3, 4, 5, 7, 8, 10, 100, 1000 – equivalence, addition, subtraction, multiplication, estimation, expressed as decimal)	✓	✓	✓
	The set Q^+: decimals, place value, addition, subtraction, multiplication, division, rounding, approximation, percentages, equivalence of fractions, decimals, percentages	✓	✓	✓
	The set R: order, every point on the number line represents a real number	✓	✓	
	Commutative property, squares and square roots, priority of operations	✓	✓	✓
	Scientific notation	✓	✓	
	Commutative and associative properties as applied to addition, subtraction, multiplication and division	✓	✓	
Applied Arithmetic & Measure	Bills; percentage profit and discount; rates and tax; VAT; compound interest, annual interest	✓	✓	✓
	Basic units of length, mass, time (including 24 hour clock and transport timetables)	✓	✓	✓
	Multiples and submultiples of these basic units	✓	✓	
	Relationship between speed, distance and time	✓	✓	✓
	Perimeter	✓	✓	✓
	Area of square, triangle and rectangle	✓	✓	✓
	Volume of rectangular solids	✓	✓	✓
	Surface area of rectangular solids	✓	✓	
	The relationship between circumference, diameter and π	✓	✓	✓
	Application of π in various formulae: $2\pi r$, πr^2, $2\pi r^2h$	✓	✓	✓
	Application of π in various formulae: $1/3\pi rh$, $4\pi r^2$, $4/3\pi r^2$	✓	✓	
	Calculating distance from a map, use of scales on drawings			✓
Functions & Graphs	Use of function notation	✓	✓	✓
	Plotting points	✓	✓	✓
	Drawing graphs of linear functions	✓	✓	✓
	Concepts of function, domain, codomain, range	✓	✓	
	Drawing graphs of quadratic functions	✓	✓	
	Minimum and maximum value of quadratic functions found graphically	✓		
	Graphing solution sets: linear inequalities in one variable	✓	✓	
	Graphical treatment of solution of first degree equations in two variables	✓	✓	
	Solution of quadratic inequality found from the graph of a quadratic function	✓		

Note. H = Higher level; O = Ordinary level, and F = Foundation level. In the case of geometry, only those theorems that are examined are listed. At Ordinary level, no theorems are examined. At Foundation level, these are presented as facts. Constructions and corollaries of theorems are not listed.
Source: Department of Education and Science (1999a, 2000b); Oldham (personal communications, 2001, 2004)

Table 6.1. Continued

Topic	*Sub-topic*	*H*	*O*	*F*
Algebra	Concepts of unknown and variable in context of formulae	✓	✓	✓
	Meaning of constant, term, expression, and coefficient	✓	✓	
	Evaluation of expressions	✓	✓	✓
	Rearrangement/simplificatin of formulae	✓	✓	✓
	Addition, subtraction, multiplication, and division	✓	✓	
	Distributive property, factors (including quadratics and squares)	✓	✓	
	First degree equations in one variable	✓	✓	✓
	First degree equations in two variables	✓	✓	
	Quadratic equations	✓	✓	
	Linear equalities in one variable	✓	✓	
Geometry	*Synthetic geometry:*			
	The plane, line segments, half line, collinear points, types of triangles	✓	✓	✓
	Naming conventions for triangles, angles, types of angles, parallel and perpendicular lines	✓	✓	✓
	Types of quadrilaterals (e.g., rhombus, parallelogram), circle, and concept of area in relation to these [area in Higher and Ordinary only]	✓	✓	✓
	Theorem: Vertically opposite angles are equal in measure	✓	✓	✓
	Theorem: Measures of three angles of a triangle sum to 180 degrees	✓	✓	✓
	Theorem: An exterior angle of a triangle equals the sum of the two interior opposite angles	✓		
	Congruency of triangles	✓	✓	
	Theorem: If two sides of a triangle are equal in measure, then the angles opposite these are equal in measure	✓	✓	
	Theorem: Opposite sides and opposite angles of a parallelogram are respectively equal in measure	✓	✓	
	Theorem: The measure of the angle at the centre of the circle is twice the measure of the angle at the circumference, standing on the same arc	✓		
	Theorem: A line through the centre of a circle perpendicular to a chord bisects that chord	✓		
	Theorem: If two triangles are equiangular, the lengths of corresponding sides are in proportion	✓		
	Theorem of Pythagoras	✓	✓	✓
	Transformation geometry:			
	Translation, central symmetry, axial symmetry.	✓	✓	✓
	Axis and centre of symmetry	✓	✓	
	Rotation	✓		
	Co-ordinate geometry:			
	Co-ordinating the plane	✓	✓	
	Co-ordinates of images under translation, axial symmetry, central symmetry	✓	✓	
	Slope of a line	✓	✓	
	Equation of a line ($y - y_1 = m(x - x_1)$)	✓	✓	
	Equation of a line ($ax + by + c = 0$; $y = mx + c$)	✓		
	Intersection of lines using algebraic methods	✓	✓	
Trigonom-etry	Cosine, sine and tangent of angles 0 to 360 degrees	✓		
	Cosine, sine and tangent of angles 0 to 90 degrees		✓	
	Functions of 30, 45, 60 degrees in surd form	✓		
	Value of angle given value of sin, cos, tan		✓	
	Solving right-angled triangle problems	✓	✓	
	Sine rule; application of formulae such as ½ ab sin c, ½ bc sin a, ½ ca sin b etc. for area.	✓		
Statistics	Collecting, recording and tabulating data	✓	✓	✓
	Bar charts, pie charts, trend graphs: drawing and interpreting	✓	✓	✓
	Discrete array expressed as a frequency table	✓	✓	✓
	Histograms: drawing and interpreting	✓		
	Mean, mode	✓	✓	✓
	Cumulative frequency	✓		
	Interquartile range, ogive, mean of grouped frequency distribution, median	✓		

Note. H = Higher level; O = Ordinary level, and F = Foundation level. In the case of geometry, only those theorems that are examined are listed. At Ordinary level, no theorems are examined. At Foundation level, these are presented as facts. Constructions and corollaries of theorems are not listed.
Source: NCCA/Department of Education and Science (2000, 2002); Oldham (personal communications, 2001, 2004)

Appendix C (Section C.3) explores further the differences between the content and style of PISA mathematics and Junior Certificate mathematics through an examination of some of the questions on the 2003 Junior Certificate mathematics examination with reference to some of the PISA mathematics items described in Appendix A. This comparison, albeit selective, illustrates some of the differences between the two assessments, and suggests that while in general Irish students have basic mathematical skills, they may lack experience in actively applying mathematical skills in novel situations, which often requires a very good conceptual understanding of the mathematics involved.

The length of the Junior Certificate mathematics examinations suggests that students would be familiar with a two-hour testing session such as PISA. Foundation level students take one two-hour paper; Ordinary level students take two two-hour papers, and Higher level students take two two-and-a-half hour papers. Both assessments also use short response formats and commonly require students to show their work or explain their reasoning; however PISA also uses multiple-choice item formats, which the Junior Certificate mathematics examination does not. There is also more emphasis in Junior Certificate mathematics on formal argumentation and proof in abstract (or 'intra-mathematical') contexts, particularly at Higher level.

The approaches to marking students' work for the two assessments are essentially different. The marking schemes for PISA treat most questions as simply right or wrong although partial credit is applied to some items; that is, three rather than two levels of marks are awarded (zero credit, some credit, full credit). The marking schemes for Junior Certificate mathematics, in contrast, are more detailed, and a zero mark for a question is less common (see http://www.examinations.ie). This is because questions are present in units, each of which is allocated a maximum mark, typically 5 or 10 marks. Each line of the student's work is scrutinised and subjected to penalties; for example, one mark may be deducted for an arithmetical slip and three for a more serious error such as a misapplication of an algebraic rule. The same error is penalised once only in any one part of a question. Furthermore, in the application of penalties, a student's mark is not allowed to drop below the 'attempt mark' for that part which is usually one third of the maximum mark for the section. Therefore, a student who makes a worthwhile attempt at a question with a maximum mark of 10 will receive at least 3 marks. These differences might suggest that the marking of mathematics in the Junior Certificate Examination offers greater scope for recognising merit in students' work than the PISA assessment.

PERFORMANCE ON THE 2003 JUNIOR CERTIFICATE MATHEMATICS, ENGLISH AND SCIENCE EXAMINATIONS

Notwithstanding differences in purpose and content between PISA and the Junior Certificate, a brief review of the performance of Irish students on the 2003 Junior Certificate Examination in mathematics, English and science are considered to place PISA results in the context of internal national standards. The data reported in this section are taken from the State Examinations Commission website (http://www.examinations.ie).

Mathematics

In total, 58441 students took mathematics at Junior Certificate level in 2003. Of these, 40.6% took the examination at Higher level, 46.9% at Ordinary level, and 12.5% at Foundation level. While almost identical percentages of males and females took the examination at Ordinary level, slightly more males (14.6%) than females (10.4%) took the examination at Foundation level, the reverse is true for Higher level (38.6% of males compared to 42.7% of females).

A breakdown of grades obtained by syllabus level and gender indicates that, at Higher and Foundation levels, just under 4% did not achieve at least a grade D. This percentage is slightly higher at Ordinary level (7.6%). The aggregate across levels (5.5%) is considerably lower than the 16.8% of Irish students who have a mean combined mathematics score in PISA 2003 that is at or below Level 1 (see Table 3.11). Gender differences in grades awarded are modest, although, as noted, slightly more females took the Higher level course, and the percent of female students obtaining grade C or above is generally a little higher than the percent of males.

Table 6.2. Breakdown of 2003 Junior Certificate Examination Mathematics Results, by Syllabus Level and Gender

		Examination Grade (Examination Mark Range)							
Level/ Gender	*Total N*	*A (85-100%)*	*B (70-84%)*	*C (55-69%)*	*D (40-54%)*	*E (25-39%)*	*F (10-24%)*	*NG (0-9%)*	*Total*
Higher									
Males	11352	16.7	31.7	29.0	18.4	3.6	0.6	0.0	100.0
Females	12382	17.7	35.2	28.3	15.7	2.6	0.4	0.0	100.0
All	23734	17.2	33.6	28.6	17.0	3.1	0.5	0.0	100.0
Ordinary									
Males	13781	7.6	28.6	32.1	22.7	6.8	2.2	0.1	100.0
Females	13602	10.8	33.5	30.5	19.0	4.9	1.3	0.1	100.0
All	27383	9.2	31.0	31.3	20.8	5.8	1.8	0.1	100.0
Foundation									
Males	4298	16.6	37.9	28.3	13.7	3.1	0.4	0.0	100.0
Females	3026	13.7	37.8	31.3	13.6	3.2	0.4	0.0	100.0
All	7324	15.4	37.8	29.5	13.6	3.2	0.4	0.0	100.0

Note. NG = No Grade. Souce: http://www.examinations.ie

Overall Strengths and Weaknesses Identified by Chief Examiners in Student Performance on the Junior Certificate Mathematics Examination

The 2003 Chief Examiners' Report on Junior Certificate Mathematics (State Examinations Commission, 2003) provides a detailed description of the performance of candidates on the examination. Many of the weaknesses (and strengths) identified are relevant in considering the performance of Irish students on PISA 2003 mathematics. The report states that many students appear to approach mathematics in a mechanical manner and are not using higher-order reasoning in working out answers, and that some fundamental conceptual understanding is lacking. Aspects of geometry, algebra and trigonometry were identified as general areas of weakness. In contrast, students typically performed well on questions that called for the application of basic concepts involving number, applied arithmetic, statistics and functions.

English

In total, 58716 students took English at Junior Certificate level in 2003. Of these, 63.1% took the examination at Higher level, 32.5% at Ordinary level, and 4.5% at Foundation level. Thus, compared to mathematics, more students take English at Higher level, and fewer take it at Foundation level. There are marked gender differences in the percentage of students taking each level. Just over half of males (56.0%), compared with 70.1% of females took the English examination at Higher level; about two-fifths of males (38.2%) compared with just over a quarter of females (26.7%) took it at Ordinary level; and almost twice as many males took the examination at Foundation level (5.8% compared to 3.1%). A breakdown of grades obtained by syllabus level and gender is shown in Table 6.3.

Not only are more females than males taking Higher level English, but more females achieve moderate to high grades. As with mathematics, the rates obtaining E, F or no grade for Junior Certificate English at all three syllabus levels are low (below 3%) and somewhat at odds with the PISA finding in both 2000 and 2003 that 11.0% of Irish students are at or below reading proficiency Level 1, the most basic level of reading literacy measured by PISA.

Table 6.3. Breakdown of 2003 Junior Certificate Examination English Results, by Syllabus Level and Gender

		Examination Grade (Examination Mark Range)							
Level/ Gender	*Total N*	*A (85-100%)*	*B (70-84%)*	*C (55-69%)*	*D (40-54%)*	*E (25-39%)*	*F (10-24%)*	*NG (0-9%)*	*Total*
Higher									
Males	16551	5.0	23.2	42.4	26.9	2.3	0.2	0.0	100.0
Females	20472	10.4	33.8	39.8	15.2	0.8	0.1	0.0	100.0
All	37023	8.0	29.1	40.9	20.4	1.5	0.1	0.0	100.0
Ordinary									
Males	11277	5.4	26.2	42.8	24.1	1.3	0.2	0.0	100.0
Females	7795	13.6	38.2	36.4	11.3	0.4	0.1	0.0	100.0
All	19072	8.8	31.1	40.2	18.8	0.9	0.1	0.0	100.0
Foundation									
Males	1702	10.2	30.8	38.7	16.5	2.9	0.8	0.1	100.0
Females	919	16.8	39.7	32.0	10.0	1.4	0.0	0.1	100.0
All	2621	12.5	33.9	36.4	14.2	2.4	0.5	0.1	100.0

Note. NG = No Grade. Souce: http://www.examinations.ie

Science

In total, 51090 students took science at Junior Certificate level in 2003; this represents 86.1% of the entire cohort (90.1% of all males and 82.0% of all females). Of those taking science, 63.9% took the examination at Higher level, and 36.1% at Ordinary level. (Unlike mathematics and English, there is no Foundation level examination for Junior Certificate science.) There are gender differences in the percentage of students taking each level: three-fifths of males (59.1%), compared with 69.3% of females took the examination at Higher level.

A breakdown of grades obtained by syllabus level and gender is shown in Table 6.4. About 5% of students at Higher level and just over 3% at Ordinary level did not achieve at least grade D. In contrast, 44.1% of students taking Higher level, and 39.1% taking Ordinary level, achieved grade A or B.

Table 6.4. Breakdown of 2003 Junior Certificate Examination Science Results, by Syllabus Level and Gender

		Examination Grade (Examination Mark Range)							
		A	*B*	*C*	*D*	*E*	*F*	*NG*	
Level/ Gender	*Total N*	*(85-100%)*	*(70-84%)*	*(55-69%)*	*(40-54%)*	*(25-39%)*	*(10-24%)*	*(0-9%)*	*Total*
Higher									
Males	15905	12.3	27.5	32.8	21.9	4.8	0.7	0.1	100.0
Females	16762	17.2	30.9	31.0	17.0	3.4	0.4	0.0	100.0
All	32667	14.8	29.3	31.9	19.4	4.0	0.5	0.0	100.0
Ordinary									
Males	11007	6.2	31.9	39.0	19.5	3.0	0.4	0.1	100.0
Females	7416	6.8	33.8	37.6	18.4	3.1	0.3	0.0	100.0
All	18423	6.4	32.7	38.4	19.0	3.0	0.4	0.0	100.0

Note. NG = No Grade. Souce: http://www.examinations.ie

RELATIONSHIP BETWEEN PERFORMANCE ON THE JUNIOR CERTIFICATE EXAMINATION AND PISA 2003

Performance on the 2003 Junior Certificate Mathematics, English and Science Converted to the Junior Certificate Performance Scale (JCPS)

To make direct comparisons between performance on PISA and performance on the Junior Certificate Examination, it is useful to put the performance of all students, regardless of the level at which they took the examination, on the same scale. Inset 6.1 describes how this scale was constructed.

Inset 6.1. Junior Certificate Performance Scale (JCPS) Scores

The 12-point Junior Certificate Performance Scale (JCPS) shown below, which has a three-grade overlap between examination levels, has been used in analyses published in the Irish national report on PISA 2000 (Shiel et al., 2001), and in a number of earlier studies (e.g., Kellaghan & Dwan, 1995; Martin & Hickey, 1993). Although such scores have in earlier studies referred to the mean performance of students on all subjects they took in the Junior Certificate Examination, they are used in analyses of the PISA data to denote performance on a single subject (mathematics, English, or science). Because there is no Foundation level science in the Junior Certificate Examination, a 9-point scale was used ranging from 4 (F, Ordinary level) to 12 (A, higher level).

Higher	**Ordinary**	**Foundation**	***JCPS 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*

Table 6.5 shows the Junior Certificate Examination achievements of the *population* of students in mathematics, English, and science in 2003, converted from raw grades to Junior Certificate Performance Scale (JCPS) scores. As the description of the performance of students on the examination in terms of their raw grades suggests, there is a small gender difference favouring females for mathematics (females score about 0.4 of a JCPS score point, or one-sixth of a standard deviation higher than males), a larger gender difference favouring females (0.9 of a JCPS score point or about half a standard deviation) for English, and a difference of just over one-quarter of a standard deviation (0.5 score points), again favouring females, for science. The magnitude of the gender difference may fluctuate somewhat from year to year. Since these data represent the population of students taking the Junior Certificate in 2003 there is no need to test formally for the statistical significance of differences. It can be concluded that gender differences are substantial for English, and smaller for mathematics and science.

Table 6.5. Percent of the Student Population Taking the Junior Certificate Mathematics, English and Science Examinations Achieving Each Level on the Junior Certificate Performance Scale (JCPS), and Mean JCPS, Overall and by Gender

JCPS	*12*	*11*	*10*	*9*	*8*	*7*	*6*	*5*	*4*	*3*	*2*	*1*	*Total*	*Mean*	*SD*
Mathematics															
Males	6.4	12.2	11.2	10.6	14.8	15.3	13.0	8.7	5.2	2.0	0.5	0.1	100.0	7.94	2.39
Females	7.6	15.0	12.1	11.8	16.8	14.5	10.3	6.2	3.9	1.4	0.3	0.0	100.0	8.33	2.31
All	7.0	13.6	11.6	11.2	15.8	14.9	11.7	7.5	4.5	1.7	0.4	0.0	100.0	8.13	2.36
English															
Males	2.8	13.0	23.8	17.2	11.3	16.4	9.8	2.3	2.3	0.9	0.2	0.0	100.0	8.57	1.98
Females	7.3	23.7	27.9	14.3	10.7	9.8	3.5	1.4	1.0	0.3	0.0	0.0	100.0	9.44	1.78
All	5.0	18.3	25.8	15.7	11.0	13.1	6.7	1.8	1.7	0.6	0.1	0.0	100.0	9.00	1.93
Science															
Males	7.3	16.3	19.4	15.5	15.9	16.3	8.0	1.2	0.2	0.0	0.0	0.0	100.0	8.95	1.82
Females	11.9	21.5	21.5	13.9	12.7	11.8	5.6	0.9	0.1	0.0	0.0	0.0	100.0	9.43	1.81
All	9.5	18.7	20.4	14.7	14.4	14.2	6.9	1.1	0.1	0.0	0.0	0.0	100.0	9.18	1.83

Note. Students attaining 'no grade' are not included in the table. For a description of how the JCPS is derived, see Inset 6.1.

The PISA 2003 cohort includes students who took the Junior Certificate Examination in 2002 (34.3% of the cohort) as well as students who took the examination in 2003 (59.6%); a further 6.1% of students took the examination in other years and are not included in these analyses. Junior Certificate results for English and/or mathematics are available for 93.9% of students for whom PISA mathematics and reading scores are available; and for science, they are available for 82.5% of students who participated in the PISA assessment of science.[48] The percentage is lower for science because, as noted, science is an optional subject. Junior Certificate results for the three subjects were placed on the 12-point Junior Certificate Performance Scale (JCPS; 9 points in the case of science, see Inset 6.1 for a description of the JCPS).

In the Irish report for PISA 2000, it was noted that students who sat Junior Certificate English and Mathematics in 1999 performed significantly better on the Junior Certificate than those who sat it in 2000, and that there was no difference in the mean scores of students taking Junior Certificate science in 1999 and 2000 (Shiel et al., 2001). Similar comparisons were made for the PISA 2003 cohort (Table 6.6) and although, across all three subject areas, students taking the examination in 2002 outperformed those taking the examination in 2003, the mean score differences are significant only for English and science; further, they are

[48] The corresponding percentages for the analyses of Junior Certificate and performance on PISA 2000 data are almost identical at 94.1% for English and mathematics and 80.5% for science.

relatively small (around one-sixth of a standard deviation), and may not be of substantive significance.[49] It should further be noted that students taking the Junior Certificate mathematics examination in 2002 did so under the old syllabus since 2003 was the first year of examination of the current syllabus. Since the primary focus here is on the PISA cohort, however, results from both years were combined into a single analysis.

Table 6.6. Mean Junior Certificate Performance Scale (JCPS) Scores in Mathematics, Science and English of Junior Certificate Examination Candidates in 2002 and 2003 – PISA 2003 Cohort

	N	*Mean*	*SE*	*Diff*	*SE Diff*	*CI 95 L*	*CI 95 U*
Mathematics							
2002	1330	8.66	0.094				
2003	2312	8.50	0.077	0.151	0.100	-0.048	0.350
All available	3642	8.56	0.068				
English							
2002	1331	9.61	0.067				
2003	2312	9.37	0.056	0.241	0.080	**0.083**	**0.400**
All available	3643	9.46	0.046				
Science							
2002	1187	9.53	0.076				
2003	2013	9.28	0.067	0.255	0.086	**0.085**	**0.426**
All available	3200	9.37	0.057				

Note. N for science is lower than those for mathematics and English since science is an optional subject. Approximately 109 PISA students (2.8%) were in Grades 7 and 8 (First and Second years) at the time of the assessment and would not have taken the JCE in 2002 or 2003.

Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Table 6.7 shows the mean Junior Certificate Performance Scale (JCPS) scores for mathematics, English and science for males and females who participated in PISA 2003. There is no significant gender difference in the mean JCPS scores for mathematics, while females significantly outperform males in science and, to an even greater extent, in English.

Table 6.7. Junior Certificate Performance Scale (JCPS) Scores for Mathematics, English and Science, by Gender: PISA 2003 Students Taking the Examinations in 2002 or 2003

	Frequencies (Maths)		*JCPS maths*		*Frequencies (English)*		*JCPS English*		*Frequencies (Science)*		*JCPS Science*	
	%T	*%A*	*Mean*	*SE*	*%T*	*%A*	*Mean*	*SE*	*%T*	*%A*	*Mean*	*SE*
Males	46.9	50.0	8.50	0.07	46.9	50.0	9.14	0.06	39.1	52.6	9.25	0.06
Females	46.9	50.0	8.63	0.10	46.9	50.0	9.82	0.06	43.4	47.4	9.51	0.09
Missing	6.1	0.0			6.1	0.0			17.5	0.0		
All Available	93.9	100.0	8.56	0.07	93.9	100.0	9.46	0.05	82.5	100.0	9.37	0.06

Mean Score Differences between Males and Females (Reference Category: Males)

	Maths				*English*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Females-Males	-0.1	0.12	-0.1	0.4	-0.7	0.09	**0.5**	**0.8**	-0.3	0.11	**0.0**	**0.5**

Note. There are no JCPS data for missing cases since these are the students who have no data for the Junior Certificate examination. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

[49] The means and standard deviations for the three JCPS scales for PISA 2003 students are as follows: mathematics mean = 8.36, SD = 2.121; English mean = 9.32, SD = 1.657; science mean = 9.37, SD = 1.726. As with the achievement data reported in Chapters 3, 4 and 7, the data in Tables 6.6 to 6.9, and 6.16 to 6.21 are weighted using the PISA student weights.

These (unadjusted) gender differences are consistent with those observed for all students taking the Junior Certificate Examination in 2003 as noted earlier (Table 6.5), but differ in a number of ways from those reported for PISA 2003 (see Chapter 4). Although in PISA 2003 reading literacy, female students outperform their male counterparts, in PISA 2003 mathematics, males significantly outperformed females, while in PISA 2003 science, there is no significant gender difference.

Table 6.8 shows the mean PISA combined mathematics, reading and science scores of students divided into high, medium and low groups based on their JCPS scores for mathematics, English and science, respectively[50]. Students in the high category of the JCPS distribution in each subject area, have, on average, corresponding PISA scores that are about 1.6 standard deviations higher than those in the low group (138 score points in mathematics and in science, and 143 points in reading). The difference between the high and medium groups and between the medium and low groups is about half this magnitude (0.8 standard deviations), regardless of subject area.

Table 6.8. Combined PISA Mathematics, Reading and Science Scores by High, Medium and Low Scores on the Junior Certificate Performance Scales (JCPS) for Mathematics, English and Science

	Frequencies (Maths)		*Maths (Comb.)*		*Frequencies (Reading)*		*Reading*		*Frequencies (Science)*		*Science*	
	%T	*%A*	*Mean*	*SE*	*%T*	*%A*	*Mean*	*SE*	*%T*	*%A*	*Mean*	*SE*
Low	34.6	36.9	438.6	2.32	26.3	28.0	442.3	3.60	26.6	32.2	447.4	8.82
Medium	28.4	30.2	510.8	2.22	44.4	47.3	529.4	2.14	30.5	37.0	523.8	5.17
High	30.9	32.9	576.9	2.36	23.2	24.7	585.3	2.67	25.4	30.8	585.7	2.61
Missing	6.1	0.0	456.3	9.48	6.1	0.0	463.5	10.79	17.5	0.0	456.2	9.22
All Available	93.9	100.0	505.9	2.45	93.9	100.0	518.9	2.66	82.5	100.0	518.3	5.03

Mean Score Differences (Reference Category: High)

	Maths (Comb.)				*Reading*				*Science*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Low-High	-138.3	3.12	**-145.9**	**-130.6**	-143.0	4.22	**-153.3**	**-132.6**	-138.4	9.14	**-160.7**	**-116.0**
Med-High	-66.1	3.06	**-73.5**	**-58.6**	-55.8	2.86	**-62.8**	**-48.8**	-62.0	5.55	**-75.6**	**-48.4**
Miss-High	-120.6	9.92	**-144.8**	**-96.3**	-121.8	11.16	**-149.1**	**-94.5**	-129.6	9.71	**-153.3**	**-105.8**

Note. N = 3880. %T = percentage all; %A = percentage available (these differ across domains since the Junior Certificate examination subject differs).

Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

An examination of correlations between students' JCPS scores and their performance on the corresponding PISA domains reveals a substantive relationship[51] in all three assessment domains (Table 6.9; see also Inset 6.2 for interpretation of the table). In all cases, correlations exceed .65. The correlation between PISA combined mathematics and the JCPS for mathematics is the highest at .75, while it is a little lower for reading and science (.67 in both cases). Among the mathematics subscales, the correlation between performance on Junior Certificate mathematics and the Space & Shape subscale is weakest at .68, while it is .73 or .74 for the other three subscales. In PISA 2000, correlations between performance on the three PISA domains and the corresponding Junior Certificate Performance Scales were similar (all in the region of .73; Shiel et al., 2001).

[50] JCPS scores were classified as 'high', 'medium' and 'low' using the same procedure as for the continuous explanatory variables reported in Chapter 4. See Chapter 4, Inset 4.2.

[51] Although the conventions used in Chapter 4 would suggest that correlations above .60 are in the strong range, since two assessments are being compared, one would expect higher correlations than those between an explanatory variable and an achievement outcome.

Overall, these findings suggests that, despite differences in context, content, and method of assessment between PISA and the Junior Certificate Examinations, which are explored further in the following section, there is an overlap in the achievements assessed by the two measures. Comparing correlations between the Junior Certificate and the corresponding PISA domains with those between the PISA 2003 domains themselves, however, one can see that students' performance on PISA areas are more highly correlated. The correlation between scores on PISA mathematics and PISA reading is .80, while for PISA mathematics and science, it is .84, and for PISA reading and PISA science, it is .85 (see Chapter 3, Table 3.31).

Inset 6.2 Interpreting Relationships Between Junior Certificate Performance Scale (JCPS) Scores and Performance on PISA

In evaluating the importance of a variable one can consider the magnitude of its regression coefficient. However, the interpretation of unstandardized coefficients can be problematic because their magnitude depends on the underlying scale of measurement. Therefore, it has become common practice to provide *standardized regression coefficients*. The standardized regression coefficients represent the predicted change in the response variable for a change of one standard deviation in an explanatory variable. It is also useful to bear in mind that, if Normally distributed, scores for two-thirds of the population lie within one standard deviation above and below the mean. One can transform raw (unstandardized) regression coefficients (see below, *a*) to describe the increase in the PISA 2003 scales associated with a one standard deviation increase in the Junior Certificate Performance Scales (*b, c*), as well as the correlation coefficient (*d*) and the associated t-statistic and its significance (*e, f*).

	(a)	*(b)*	*(c)= (a)*(b)*	*(d)*	*(e)*	*(f)*
	Raw Coefficient	*SD JCPS Maths*	*SD Unit*	*r*	*t*	*p*
PISA comb. maths	29.67	2.121	62.94	.754	49.944	<.001

Table 6.9. Associations Between Performance on PISA 2003 Combined Mathematics and Mathematics Subscales and Junior Certificate Mathematics, PISA 2003 Reading and Junior Certificate English, and PISA 2003 Science and Junior Certificate Science

Comparison	*Raw Coeff*	*SE*	*SD Unit*	*r*	*t*	*p*
PISA maths (combined) - JCE maths	29.67	0.59	62.94	**.754**	49.94	<.001
PISA maths (space & shape) - JCE maths	29.77	0.71	63.13	**.681**	41.67	<.001
PISA maths (change & relationships) - JCE	29.90	0.56	63.41	**.741**	53.04	<.001
PISA maths (uncertainty) - JCE maths	30.46	0.58	64.60	**.743**	52.18	<.001
PISA maths (quantity) - JCE maths	29.84	0.60	63.28	**.731**	49.84	<.001
PISA reading - JCE	34.23	0.86	56.71	**.673**	39.99	<.001
PISA science - JCE science	32.99	0.89	56.94	**.674**	39.05	<.001

Note. The column 'SD Unit' is obtained by by multiplying the raw coefficienct by the standard deviation associated with the Junior Certificate Performance Scale in question (SD for mathematics = 2.121, English = 1.657, science = 1.726). Significant correlations (p ≤ .05) are highlighted in bold.

THE PISA 2003 TEST-CURRICULUM RATING PROJECT

This section describes the outcomes of the test-curriculum rating project for the PISA 2003 mathematics items. As noted in Chapters 1 and 2, PISA departs from more traditional curriculum-based assessments in that it is designed to assess real-life literacy skills deemed relevant for adult life. Further, as noted earlier in this chapter, the PISA assessment of mathematics represents a departure from the Junior Certificate mathematics examination. A comparison of the content of the two assessments is of interest in interpreting strengths and weaknesses in performance on PISA and also in informing discussion and debate about the mathematics curriculum in Ireland. Chapter 2 described in some detail the rationale underlying the test-curriculum rating project as it was implemented in 2000, and the outcomes for mathematics for PISA 2000. This section focuses on the outcomes for 2003.

The framework for the 2003 test-curriculum rating project is identical to that used in 2000. The framework focuses on the expected familiarity of Third year (Grade 9) students with each PISA mathematics item, distinguishing between syllabus level (Higher, Ordinary, Foundation) and item aspect (concept, context of application, and format). Appendix C (Section C.2) describes these in more detail. Each item receives nine ratings, one for each aspect at each syllabus level, with ratings ranging from 1 (‘not familiar’) to 3 (‘very familiar’).

It was noted in Chapter 3 that the performance of Irish students in PISA 2003 mathematics was around the OECD average, and that Irish students displayed quite an uneven profile of performance in mathematical content areas: they performed significantly below the OECD mean on the Space & Shape subscale, at the OECD mean on the Quantity subscale, and significantly above the OECD mean on both the Change & Relationships and the Uncertainty subscales. Thus an examination of the expected familiarity with items by subscale is of interest in exploring these differences with respect to the national curriculum (i.e., the Junior Certificate mathematics syllabus and examination).

For those items which were deemed to be somewhat or very familiar in terms of the concepts they tapped, the Junior Certificate area into which the item fell was identified. The revised syllabus and teaching guides, as well as Junior Certificate mathematics examination papers[52] were used as reference documents in this exercise (Department of Education and Science/ National Council for Curriculum and Assessment, 2000, 2002). In addition to familiarity ratings, items were rated in terms of whether a calculator would be useful or not.

In the case of the PISA 2000 mathematics items which were retained in the PISA 2003 assessment to measure change over time (‘link items’), the same rating was again applied in 2003, but ratings for link items were reviewed in the event that a change in the current mathematics syllabus might result in a change in the 2000 rating.[53]

In 2000, the test-curriculum rating project was also carried out for reading and mathematics, but these were not repeated for 2003 since all PISA 2003 reading items were taken from the PISA 2000 reading item pool, and the majority of science items were taken from the PISA 2000 science item pool. Further, there were no revisions to the syllabuses in the corresponding Junior Certificate subjects. In PISA 2006, when science becomes the major assessment domain, it is envisaged that the test-curriculum rating project will focus on science. The rating in 2003 was carried out by three individuals with extensive expertise in the Irish mathematics curricula and assessments.[54] The following section describes outcomes of the PISA 2003 test-curriculum rating project at the item level, while a subsequent section examines student-level performance in terms of the test-curriculum rating outcomes.

[52] Examination papers for 2003 are available at http://www.examinations.ie
[53] In practice, the ratings for the common items remained the same in all cases.
[54] The same three individuals also carried out the ratings in 2000.

Initially, ratings were made independently, and, following a meeting of all raters[55], consensus on the final set was obtained. 'Consensus' was defined as the modal rating assigned to a particular scale at a particular syllabus level where there was either perfect agreement across the three raters *or* where ratings did not differ by more than one scale point. Thus, items which had received ratings on all three points of the scale were automatically flagged, as were items on which ratings differed by more than one scale point across raters. Flagged items were discussed at the consensus meeting and a final rating agreed upon. (Table C.1 in Appendix C, Section C.2 shows the percentage of items in each of the three domains on which there was a lack of consensus.)

Outcomes from the PISA 2003 Test-Curriculum Rating Project for Mathematics

Location of PISA 2003 Items in Junior Certificate Topic Areas

Table 6.10 shows a cross-tabulation of PISA 2003 mathematics items by Junior Certificate mathematics topic area and syllabus level. At Higher level, 28.6% of items were rated as being 'not familiar' in terms of the underlying concept and therefore not located anywhere on the Junior Certificate mathematics syllabus. Corresponding figures for Ordinary and Foundation are 33.0% and 49.4%, respectively. The Junior Certificate mathematics topic areas of sets, geometry, and trigonometry are not assessed at all by the PISA mathematics items, as evidenced in the high number of cells with zeroes in the lower half of the table. There is also little coverage by PISA mathematics of algebra or of functions and graphs. Table 6.10 also suggests that the bulk of PISA mathematics items which *are* somewhat familiar to Irish students are located in the Junior Certificate mathematics topic areas of applied arithmetic and measure, and statistics.

Table 6.10. Curriculum Area Ratings for PISA 2003 Mathematics Items Cross-tabulated with Junior Certificate Mathematics Syllabus Level

	Junior Certificate mathematics strand area									
	Not in Junior Cert.		*Number systems*		*Applied arith. & measure*		*Algebra*		*Statistics*	
Syllabus Level	N	%	N	%	N	%	N	%	N	%
Higher	26	28.6	8	8.8	30	33.0	5	5.5	18	19.8
Ordinary	30	33.0	9	9.9	29	31.9	4	4.4	16	17.6
Foundation	44	49.4	8	9.0	23	25.8	1	1.1	13	14.6
	Functions and graphs		*Sets*		*Geometry*		*Trigo-nometry*		*Total*	
Syllabus Level	N	%	N	%	N	%	N	%	N	%
Higher	4	4.4	0	0.0	0	0.0	0	0.0	91	100.0
Ordinary	3	3.3	0	0.0	0	0.0	0	0.0	91	100.0
Foundation	0	0.0	0	0.0	0	0.0	n/a	n/a	89	100.0

Note. Total number of PISA 2003 mathematics items = 85. As evidenced in the totals, 6 items were identified as being located in two Junior Certificate strand areas in the case of higher and ordinary levels, and 4 items in the case of foundation level.

In all items with two strand areas identified, raters felt that students drew on both number systems and applied arithmetic and measure. Trigonometry is not covered in the foundation level syllabus.

[55] An independent observer was present at this meeting to help ensure that the final set of ratings for each item was reached in an objective manner.

Table 6.11 shows a cross-tabulation of PISA 2003 subscale areas (Space & Shape, Change & Relationships, Quantity, and Uncertainty) with the topic areas identified in the Junior Certificate mathematics syllabus by level.

Table 6.11. *Curriculum Area Ratings Cross-tabulated with PISA 2003 Mathematics Items (Subscales), by Junior Certificate Syllabus Level*

	PISA 2003 Mathematics Subscale							
Area of Junior Cert./Level	*Change and Relationships (N items = 22)*		*Quantity (N items = 23)*		*Space and Shape (N items = 20)*		*Uncertainty (N items = 20)*	
Higher Level	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Not in Junior Cert.	5	21.7	6	21.4	6	30.0	9	45.0
Number systems	1	4.3	7	25.0	0	0.0	0	0.0
Applied arith. & measure	3	13.0	13	46.4	13	65.0	1	5.0
Algebra	5	21.7	0	0.0	0	0.0	0	0.0
Statistics	6	26.1	2	7.1	0	0.0	10	50.0
Functions and graphs	3	13.0	0	0.0	1	5.0	0	0.0
Sets	0	0.0	0	0.0	0	0.0	0	0.0
Geometry	0	0.0	0	0.0	0	0.0	0	0.0
Trigonometry	0	0.0	0	0.0	0	0.0	0	0.0
Total	23	100.0	28	100.0	20	100.0	20	100.0
Ordinary Level								
Not in Junior Cert.	6	26.1	6	21.4	7	35.0	11	55.0
Number systems	1	4.3	8	28.6	0	0.0	0	0.0
Applied arith. & measure	3	13.0	12	42.9	13	65.0	1	5.0
Algebra	4	17.4	0	0.0	0	0.0	0	0.0
Statistics	6	26.1	2	7.1	0	0.0	8	40.0
Functions and graphs	3	13.0	0	0.0	0	0.0	0	0.0
Sets	0	0.0	0	0.0	0	0.0	0	0.0
Geometry	0	0.0	0	0.0	0	0.0	0	0.0
Trigonometry	0	0.0	0	0.0	0	0.0	0	0.0
Total	23	100.0	28	100.0	20	100.0	20	100.0
Foundation Level								
Not in Junior Cert.	11	50.0	9	33.3	10	50.0	14	70.0
Number systems	0	0.0	8	29.6	0	0.0	0	0.0
Applied arith. & measure	5	22.7	8	29.6	10	50.0	0	0.0
Algebra	1	4.5	0	0.0	0	0.0	0	0.0
Statistics	5	22.7	2	7.4	0	0.0	30	30.0
Functions and graphs	0	0.0	0	0.0	0	0.0	0	0.0
Sets	0	0.0	0	0.0	0	0.0	0	0.0
Geometry	0	0.0	0	0.0	0	0.0	0	0.0
Total	22	100.0	27	100.0	20	100.0	20	100.0

Note. Total number of PISA 2003 mathematics items = 85. Percentages are column %.
As evidenced in the totals, 6 items were identified as being located in two Junior Certificate strand areas in the case of higher and ordinary levels, and 4 items in the case of foundation level.
In all items with two strand areas identified, raters felt that students drew on both number systems and applied arithmetic and measure. Trigonometry is not covered in the foundation level syllabus.

The table shows that concepts associated with each PISA subscale are, at times, distributed across several Junior Certificate mathematics areas. For example, at Higher level, concepts underlying the 18 items associated with the PISA Change & Relationships subscale whose concepts *are* on the Junior Certificate syllabus are spread across five Junior Certificate topic areas (number systems, applied arithmetic and measure, algebra, statistics, and functions and graphs). Items associated with PISA Quantity are spread across three

Junior Certificate areas (number systems, applied arithmetic and measure, and functions and graphs), although concentrated in applied arithmetic and measure. The PISA Space & Shape items rated as somewhat or very familiar in terms of their underlying concept, which one might expect to be associated with the topic area of geometry, were almost exclusively located in the Junior Certificate topic area of applied arithmetic and measure.[56] Almost all PISA Uncertainty items rated as somewhat/very familiar were located in the Junior Certificate topic area of statistics.

Table 6.12 shows a cross-tabulation of Junior Certificate mathematics areas with the PISA mathematics strand areas as described in the PISA 2003 assessment framework (OECD, 2003), and are more closely aligned to traditional curriculum areas.

Table 6.12. *Curriculum Area Ratings Cross-tabulated with PISA 2003 Mathematics Items (Strand Areas), by Junior Certificate Syllabus Level*

	PISA 2003 Maths Strand													
	Algebra (N items = 3)		*Discrete Maths (N items = 5)*		*Functions (N items = 9)*		*Geometry (N items = 18)*		*Number (N items = 27)*		*Probability (N items = 5)*		*Statistics (N items = 18)*	
Area of Junior Cert./Level	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Higher Level														
Not in Junior Cert.	1	33.3	4	80.0	3	33.3	6	33.3	2	6.1	5	100	5	27.8
Number systems	0	0.0	0	0.0	0	0.0	0	0.0	8	24.2	0	0.0	0	0.0
Applied arith. & measure	0	0.0	1	20.0	0	0.0	11	61.1	18	54.5	0	0.0	0	0.0
Algebra	2	66.7	0	0.0	2	22.2	0	0.0	1	3.0	0	0.0	0	0.0
Statistics	0	0.0	0	0.0	1	11.1	0	0.0	4	12.1	0	0.0	13	72.2
Functions and graphs	0	0.0	0	0.0	3	33.3	1	5.6	0	0.0	0	0.0	0	0.0
Sets	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Geometry	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Trigonometry	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Total	3	100.0	5	100.0	9	100.0	18	100.0	33	100.0	5	100.0	18	100.0
Ordinary Level														
Not in Junior Cert.	2	66.7	4	80.0	3	33.3	7	38.9	2	6.1	5	100	7	38.9
Number systems	0	0.0	0	0.0	0	0.0	0	0.0	9	27.3	0	0.0	0	0.0
Applied arith. & measure	0	0.0	1	20.0	0	0.0	11	61.1	17	51.5	0	0.0	0	0.0
Algebra	1	33.3	0	0.0	2	22.2	0	0.0	1	3.0	0	0.0	0	0.0
Statistics	0	0.0	0	0.0	1	11.1	0	0.0	4	12.1	0	0.0	11	61.1
Functions and graphs	0	0.0	0	0.0	3	33.3	0	0.0	0	0.0	0	0.0	0	0.0
Sets	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Geometry	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Trigonometry	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Total	3	100.0	5	100.0	9	100.0	18	100.0	33	100.0	5	100.0	18	100.0
Foundation Level														
Not in Junior Cert.	3	100	5	100	6	66.7	10	55.6	6	19.4	5	100	9	50.0
Number systems	0	0.0	0	0.0	0	0.0	0	0.0	8	25.8	0	0.0	0	0.0
Applied arith. & measure	0	0.0	0	0.0	3	33.3	8	44.4	12	38.7	0	0.0	0	0.0
Algebra	0	0.0	0	0.0	0	0.0	0	0.0	1	3.2	0	0.0	0	0.0
Statistics	0	0.0	0	0.0	0	0.0	0	0.0	4	12.9	0	0.0	9	50.0
Sets	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Geometry	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Total	3	100.0	5	100.0	9	100.0	18	100.0	31	100.0	5	100.0	18	100.0

Note. Total number of PISA 2003 mathematics items = 85. Percentages are column %.
As evidenced in the totals, 6 items were identified as being located in two Junior Certificate strand areas in the case of higher and ordinary levels, and 4 items in the case of foundation level.
In all items with two strand areas identified, raters felt that students drew on both number systems and applied arithmetic and measure. Trigonometry is not covered in the foundation level syllabus.

[56] In fact, a number of the PISA Space & Shape items involve the manipulation of representations of two- and three-dimensional objects. This is an area not strongly represented on the Junior Certificate mathematics course.

Taking Higher level as an example, one can see (top row) that the two PISA strand areas least covered at that level in the Junior Certificate syllabus are probability (concepts underlying 100% of these items are not on the Junior Certificate syllabus) and discrete mathematics (80.0% are not on Junior Certificate syllabus). In contrast, concepts underlying almost all PISA 'number' items (93.9%) are on the Junior Certificate syllabus. Looking down the columns for each PISA strand area, one can also see where in the Junior Certificate syllabus each PISA strand area is located. Some areas show a close match. For example, the PISA algebra and statistics items whose concepts were rated as familiar are located in the corresponding topic areas on the Junior Certificate syllabus. Others are more dispersed (e.g., familiar PISA functions items are located in functions and graphs but also in algebra and statistics). Broadly speaking, the pattern is the same across syllabus levels, although comparatively more PISA algebra items are not covered in the Ordinary level syllabus, and none at all at Foundation level. In addition to PISA algebra, Foundation level students would not be expected to be familiar with any of the items in the PISA discrete mathematics and probability strands (see Table 1.3, Chapter 1, for information how the PISA strand areas and PISA subscales relate to one another).

Item-Level Curriculum Familiarity Ratings

In PISA 2000, just two of the four major areas associated with mathematics in PISA 2003 were assessed. The broadening of the assessment domain since 2000, therefore, might be associated with a different pattern of curriculum familiarity ratings than that reported in PISA 2000, which in turn may have implications for the interpretation of student performance. Table 6.13 shows the curriculum familiarity ratings for PISA 2003 mathematics for each of the three item aspects (concept, context of application, and format) for each syllabus level.

Table 6.13. PISA 2003 Mathematics Curriculum Familiarity Ratings, 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

As suggested in the curriculum area ratings discussed previously (Table 6.10), the concepts underlying the majority of items at Higher (69.4%) and Ordinary (64.7%) levels were somewhat or very familiar, while just under half of the items at Foundation level (48.3%) were rated in this way. In contrast, the contexts in which the mathematics problems were presented were rated as unfamiliar for the majority of items at all three syllabus levels (65.9% at Higher level, 70.6% at Ordinary level, and 80.0% at Foundation level). Item formats were also largely unfamiliar to Irish students, regardless of syllabus level (at least in the context of Junior Certificate mathematics). In the case of both concept familiarity and the context of application, the percentages of items rated as not familiar are almost identical for all three syllabus levels when compared with the ratings for PISA 2000. This is also the case with item format, except for Foundation level, where proportionately more items were rated as unfamiliar in 2003 (see Chapter 2, Table 2.3).

To explore curriculum familiarity further, overall concept familiarity ratings were compared both for the PISA subscale areas (Table 6.14) and for the three PISA competency clusters (Table 6.15) (see Chapter 1 for a detailed description of the competency clusters). Looking first at Table 6.14, one can see that, regardless of the subscale, the percentage of items rated as unfamiliar increases as one moves from Higher to Foundation level. According to these ratings, Irish students were expected to be familiar with the concepts underlying a substantial portion of the items on the Quantity subscale (74% at Higher and Ordinary levels and 41% at Foundation level), as well as on the Change & Relationships subscale (77% at Higher, 73% at Ordinary, and 50% at Foundation level). Ratings on the Space & Shape items suggest moderate familiarity, while students were expected to be least familiar with items on the Uncertainty subscale.

Table 6.14. Curriculum Familiarity Ratings for Concept, by Junior Certificate Syllabus Level and PISA Subscale

	Not familiar		*Somewhat familiar*		*Very familiar*		*Total*	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Space and Shape								
Higher	6	30.0	6	30.0	8	40.0	20	100.0
Ordinary	7	35.0	8	40.0	5	25.0	20	100.0
Foundation	10	50.0	7	35.0	3	15.0	20	100.0
Change and Relationships								
Higher	5	22.7	8	36.4	9	40.9	22	100.0
Ordinary	6	27.3	9	40.9	7	31.8	22	100.0
Foundation	11	50.0	6	27.3	5	22.7	22	100.0
Quantity								
Higher	6	26.1	4	17.4	13	56.5	23	100.0
Ordinary	6	26.1	6	26.1	11	47.8	23	100.0
Foundation	9	39.1	7	30.4	7	30.4	23	100.0
Uncertainty								
Higher	9	45.0	3	15.0	8	40.0	20	100.0
Ordinary	11	55.0	2	10.0	7	35.0	20	100.0
Foundation	14	70.0	2	10.0	4	20.0	20	100.0

One might expect that this pattern of ratings might be reflected in students' actual achievements on the four subscales, but this is not the case. It has already been noted (Chapter 3) that Irish students scored significantly below the OECD average on the Space & Shape subscale (Table 3.3), significantly above the OECD average on the Change & Relationships (Table 3.5) and Uncertainty subscales (Table 3.7), and around the OECD average on the Quantity subscale (Table 3.9).

Looking at Table 6.15, one can again see a general pattern of decreasing familiarity with mathematics processes as one moves from Higher to Ordinary to Foundation level. The majority of all Reproduction items are expected to be somewhat or very familiar to students at all syllabus levels, while ratings on the Connections items suggest moderate familiarity. Reflection items are somewhat less familiar to students, particularly at Foundation level, where 73.7% of such items were rated as being unfamiliar.

Table 6.15. Curriculum Familiarity Ratings for Concept, by Junior Certificate Syllabus Level and PISA Process

	Not familiar		*Somewhat familiar*		*Very familiar*		*Total*	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Reproduction								
Higher	5	19.2	3	11.5	18	69.2	26	100.0
Ordinary	6	23.1	5	19.2	15	57.7	26	100.0
Foundation	10	38.5	4	15.4	12	46.2	26	100.0
Connections								
Higher	14	35.0	11	27.5	15	37.5	40	100.0
Ordinary	15	37.5	12	30.0	13	32.5	40	100.0
Foundation	20	50.0	14	35.0	6	15.0	40	100.0
Reflection								
Higher	7	36.8	7	36.8	5	26.3	19	100.0
Ordinary	9	47.4	8	42.1	2	10.5	19	100.0
Foundation	14	73.7	4	21.1	1	5.3	19	100.0

Calculator Neutrality of PISA 2003 Items

An examination of the calculator neutrality ratings suggests that the PISA 2003 assessment is not biased in a significant way in favour of students who use calculators during the assessment. The majority (70.6%) of items were rated as not needing a calculator to solve the problem; it was judged that a calculator may assist somewhat in the solution of a further 28.2% of items; and in the case of just one item it was judged that a calculator would be an advantage. Although a small but significant performance advantage for students who used a calculator during the assessment compared to those who did not was noted in Chapter 4 (see Table 4.22), other factors (e.g. those relating to problem-solving strategies used by the students) may have come into play in this relationship.

Student-Level Analyses

The PISA 2003 test design consisted of 13 test booklets, split into four half-hour blocks of test items. Each block consisted of items from one domain only (e.g., all mathematics). There were seven mathematics blocks, and two each of reading, science, and problem solving (see Appendix A, Table A.1). Every student, regardless of which booklet he/she attempted, completed at least one mathematics block. Taking into account both the mathematics syllabus level at which the student had taken/was to take the Junior Certificate mathematics examination, as well as the particular combination of mathematics items in each booklet, a booklet-level familiarity rating was assigned to each student for each of three aspects under consideration (concept, context, and format). Relationships between student performance on the PISA mathematics assessment, in the context of their expected familiarity with the particular set of items corresponding to the mathematics syllabus level that they had taken/were to take, were then examined. It should be noted that, regardless of whether a student skipped or did not reach certain items, each student taking the same booklet and at the same syllabus level is assigned the same curriculum scores for the three aspects.

Before these are described, it is relevant to point out that the scales for the three aspects are empirically as well as logically interdependent. Table 6.16 shows the intercorrelations between the curriculum familiarity scales at the student level. They are all in excess of .60 and indicate that the relationship of one scale with achievement should not be considered in isolation from the others.

Table 6.16. Linear Associations Between Concept, Context and Format Curriculum Familiarity Ratings (Mathematics)

	r	*t*	*p*
Concept-Context	**.670**	62.39	<.001
Concept-Format	**.792**	79.51	<.001
Context-Format	**.611**	62.19	<.001

Note. Significant correlations (p ≤ .05) are highlighted in bold.

Given the inter-relatedness of the scales, a principal components analysis was carried out, the results of which suggested that the scales were indeed tapping a single underlying factor which explained about 79% in the total variation of the scales.[57] Hence, a single global curriculum familiarity scale, incorporating all three aspects, was created.

Table 6.17 shows the correlations between individual curriculum aspect scales, the global scale, and achievement on the combined PISA mathematics scale (see Inset 6.2 for information on how to interpret the table). The correlation between concept familiarity and mathematics achievement, at .37, is the highest (and in the moderate range); item format correlates .28 with achievement, and context correlates .21. (The respective correlations from PISA 2000 mathematics are .48, .23 and .20.) The global curriculum familiarity scale has a moderate correlation of .32 with combined mathematics achievement. As was the case in PISA 2000, these findings suggest that concept familiarity is most strongly predictive of success on an item.

Table 6.17. Linear Associations Between Concept, Context, Format and Global Curriculum Familiarity Ratings, and Achievement on the Combined Mathematics Scale

	Combined Mathematics					
	Raw Coeff	*SE*	*SD Unit*	*r*	*t*	*p*
Concept	168.72	7.213	31.37	**.370**	23.39	<.001
Context	106.80	10.659	17.70	**.209**	10.02	<.001
Format	110.29	6.996	23.68	**.280**	15.76	<.001
Global	27.40	1.526	27.40	**.324**	17.96	<.001

Note. The column 'SD Unit' shows the increase in the outcome variable associated with a one standard deviation increase in the explanatory variable. Significant correlations (p ≤ .05) are highlighted in bold.

PERFORMANCE ON MATHEMATICS SUBSCALES BY JUNIOR CERTIFICATE MATHEMATICS SYLLABUS LEVEL

Chapter 4 showed that there were substantial differences in performance on the PISA 2003 combined mathematics scale across syllabus levels.[58] Table 6.18 (a and b) shows the mean performance by syllabus level for each of the four mathematics subscales. The performance difference is quite consistent across levels: between Higher and Foundation levels it is about two standard deviations, and between Ordinary and Foundation levels, it is about one standard deviation. The performance difference between Higher and Ordinary levels is smallest for the Space & Shape subscale (174.8 score points) and largest for the Change & Relationships subscale (186.8 score points). The performance difference also

[57] Factor loadings were .925, .846, and .902 for concept, context and format, respectively.

[58] As with analyses reported in Chapter 4, these were the actual syllabus levels taken from the Junior Certificate Examinations Database in approximately 94% of cases; in the remainder of cases, the level indicated by the student was taken from the Student Questionnaire, where available.

tends to be slightly smaller when one compares Ordinary and Foundation as opposed to Higher and Ordinary.

Table 6.18a. Mean Scores on the Space and Shape, and Change and Relationships Subscales, by Level of Junior Certificate Mathematics Syllabus Studied (Higher, Ordinary and Foundation)

	Space and Shape				*Change and Relationships*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*
Higher	42.5	42.8	536.6	2.27	566.6	1.97
Ordinary	50.0	50.3	441.8	2.25	473.1	2.08
Foundation	6.8	6.9	362.0	5.61	379.8	5.68
Missing	0.6	0.0	370.9	27.59	387.6	23.07
All Avail	100.0	99.4	476.9	2.43	506.7	2.45

Mean Score Differences (Reference Category: Ordinary)

	Space and Shape				*Change and Relationships*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Found-Ord	-79.8	5.9	**-91.6**	**-68.1**	-93.3	6.0	**-198.4**	**-175.2**
High-Ord	94.8	3.1	**-120.2**	**-63.9**	93.5	2.7	**88.1**	**98.8**
Miss-Ord	-70.9	27.4	**-125.4**	**-16.4**	-85.5	23.3	**-131.8**	**-39.2**

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Table 6.18b. Mean Scores on the Quantity, and Uncertainty Subscales, by Level of Junior Certificate Mathematics Syllabus Studied (Higher, Ordinary and Foundation)

	Quantity				*Uncertainty*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*
Higher	42.5	42.8	562.6	2.02	579.5	2.17
Ordinary	50.0	50.3	467.6	2.21	481.7	2.31
Foundation	6.8	6.9	381.4	5.82	401.3	5.84
Missing	0.6	0.0	400.2	22.09	381.3	20.34
All Avail	100.0	99.4	502.3	2.48	518.0	2.66

Mean Score Differences (Reference Category: Ordinary)

	Uncertainty				*Quantity*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Found-Ord	-80.4	6.1	**-190.1**	**-166.4**	-86.1	6.0	**-98.0**	**-74.3**
High-Ord	97.8	2.9	**92.1**	**103.5**	95.1	2.7	**89.7**	**100.4**
Miss-Ord	-100.4	20.5	**-141.2**	**-59.6**	-67.4	22.1	**-111.4**	**-23.4**

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

RELATIONSHIP BETWEEN PISA PROFICIENCY LEVELS AND JUNIOR CERTIFICATE SYLLABUS LEVELS

Referring to the cut-points associated with the PISA mathematics proficiency levels described in Chapter 3 (i.e., the points on the continuous achievement scale used to split the scale into categories), the mean score of students taking Foundation level mathematics is within Level 1 for all subscales (ranging from towards the bottom of Level 1 for the Space & Shape subscale, to about one-third from the top of Level 1 for the Uncertainty subscale); they are all within Level 2 for Ordinary level students (ranging from about one-third from the bottom of Level 2 for the Space & Shape subscale to near the top of Level 2 for the Uncertainty subscale), while the mean score of Higher level students is about one-third from the top of Level 4 for the Space & Shape subscale, and within Level 5 for the other three subscales.

Table 6.18 thus illustrates the wide variation in achievement across syllabus levels, as well as variation in achievement within each syllabus level across the three mathematics subscales, which is consistent with the variability in performance of Irish students as a whole across the subscales noted in Chapter 3 (see Table 3.18). However, it should be borne in mind that the overall variation in student achievement in Ireland is smaller than the OECD average, both for combined mathematics, and for the mathematics subscales (as already noted in Chapter 3).

PISA 2003 Mathematics Proficiency Levels Cross-tabulated with Junior Certificate Mathematics Syllabus Level

Table 6.19 shows the percentage of students at each PISA combined mathematics proficiency level, cross-tabulated with Junior Certificate syllabus level. The table shows a marked difference in the levels of proficiency demonstrated by students taking Junior Certificate mathematics at the three syllabus levels. For example, one third of Foundation level students have a mean score below Level 1; i.e., they did not demonstrate even the most basic mathematical proficiencies associated with PISA mathematics. A further two-fifths of this group (38.5%) score at Level 1, and 22.5% at Level 2. Just under 6% of students taking Foundation level surpass Level 2. No student taking Foundation level demonstrated a proficiency higher than Level 3. At Ordinary level, 21.9% of students are at or below Level 1, and close to two-fifths (36.2%) are at Level 2. Only around two-fifths of Ordinary level students (41.9%) have a mean score at Level 3 or higher, and less than 2% have a mean score at Levels 5 or 6. At Higher level, few students score at or below Level 1 (1.5%), although 10.5% are at or below Level 2. Close to one-quarter have a mean score at the highest proficiency levels (5 or 6).

Table 6.19. Percent of Irish Students at Each Combined Mathematics Proficiency Level Cross-tabulated with Junior Certificate Mathematics Syllabus Level

	Below Level 1		*Level 1*		*Level 2*		*Level 3*		*Level 4*		*Level 5*		*Level 6*	
	%	*SE*	*%*	*SE*	*%*	*SE*	*%*	*SE*	*%*	*SE*	*%*	*SE*	*%*	*SE*
Higher	0.3	0.16	1.2	0.33	9.0	1.09	28.8	1.26	35.8	1.52	19.7	1.39	5.2	0.76
Ordinary	4.1	0.70	17.8	1.26	36.2	1.27	30.4	1.30	9.9	0.90	1.5	0.38	0.1	0.11
Foundation	33.4	4.05	38.5	4.18	22.5	3.35	5.5	1.83	0.0	0.00	0.0	0.00	0.0	0.00

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

PISA 2000 and 2003 Reading Proficiency Levels Cross-tabulated with Junior Certificate English Syllabus Level

Table 6.20 shows the percentage of students at each PISA combined reading proficiency level, cross-tabulated with Junior Certificate syllabus level for 2000 and 2003. It is important to note that these proficiency levels are not comparable to those for mathematics since different criteria were used to define the proficiencies at each level, and the width of the proficiency levels is narrower for mathematics.

There are large differences in the distribution of students across proficiency levels, depending on the syllabus level taken for Junior Certificate English. For example, just under 78% of Foundation level students score at or below Level 1 (i.e. at or below the most basic level of reading measured by PISA) in 2003, and none scores above Level 3. At Ordinary level, over one-quarter of students score at or below Level 1, and just 7.4% demonstrate the more advanced reading skills associated with Levels 4 and 5. In contrast, almost half of Higher level students (48.3%) scored at Levels 4 or 5, and very few (just over 2%) at or below Level 1.

There are some differences in the distribution of performance across reading proficiency levels when one compares these for each syllabus level for 2000 and 2003. For example, while the low reading standards of Foundation level students in 2003 remain a cause for some concern, fewer Foundation level students in 2003 (77.5%) than in 2000 (90.0%) achieved reading scores at or below Level 1. However, the large standard errors associated with the mean scores and score differences at Foundation level means that obtained differences are not statistically significant. In contrast, a lower percentage of Higher level students in 2003 (13.2%) compared to 2000 (20.0%) scored at Level 5, a pattern that is consistent with the decrease in achievement in reading literacy in 2003 at the upper end of the achievement distribution noted in Chapter 3 (Table 3.29). There is little change in the distribution for Ordinary level between 2000 and 2003.

Table 6.20. Percent of Irish Students at Each Reading Proficiency Level Cross-tabulated with Junior Certificate English Syllabus Level, 2000 and 2003

	Below Level 1		*Level 1*		*Level 2*		*Level 3*		*Level 4*		*Level 5*	
2000	*%*	*SE*	*%*	*SE*	*%*	*SE*	*%*	*SE*	*%*	*SE*	*%*	*SE*
Higher	0.4	0.15	1.8	0.36	11.0	0.91	31.5	1.15	35.3	1.27	20.0	1.00
Ordinary	7.4	1.24	21.3	2.10	34.8	2.03	26.6	2.42	8.7	1.18	1.1	1.43
Foundation	51.0	9.54	39.0	8.26	7.8	6.22	2.3	2.85	0.0	0.00	0.0	0.00
2003												
Higher	0.2	0.11	2.1	0.41	13.9	1.29	35.3	1.58	35.2	1.60	13.2	0.94
Ordinary	6.0	1.12	21.1	1.81	38.5	2.15	27.0	2.05	6.6	1.04	0.7	0.27
Foundation	39.1	10.87	38.4	10.15	16.5	5.10	6.0	3.52	0.0	0.00	0.0	0.00

Note. Total number of Irish students (2003) = 3880. N Higher = 2668, N Ordinary = 1131, N Foundation = 56, N Missing syllabus level = 26.

Total number of Irish students (2000) = 3854. N Higher = 2853, N Ordinary = 1173, N Foundation = 57, N Missing syllabus level = 39.

DIFFERENCES IN MATHEMATICS ACHIEVEMENT AND ATTITUDES TOWARDS MATHEMATICS ACROSS SYLLABUS LEVELS

Performance differences on PISA mathematics associated with Junior Certificate mathematics syllabus level were noted above. Given that the current Junior Certificate mathematics syllabus has among its aims the fostering of interest and appreciation in mathematics, and the development of confidence of students' own mathematics abilities, this section looks at these concepts (as defined and measured in PISA 2003) as outcomes. PISA, as will be recalled, collected information on a number of aspects of students' interests and attitudes, including interest in mathematics and anxiety towards mathematics. Composite scales with an OECD mean of zero and standard deviation of 1 were constructed to report outcomes on these measures (see Chapter 4, Inset 4.2 for a description, and Table 4.25 and 4.26 for outcomes).

Table 6.21 compares the mean scores on the interest in mathematics and anxiety towards mathematics composite variables for each Junior Certificate mathematics syllabus level. Taking interest in mathematics first, it can be seen that Ordinary and Foundation level students have similarly low interest in mathematics, around one-quarter of a standard deviation below the OECD mean, and Higher level students have a mean interest in mathematics about one-fifth of a standard deviation above the OECD mean, and around two-fifths to one-half of a standard deviation above that of Ordinary and Foundation level students. Mathematics anxiety (where low negative values indicate low anxiety) shows a more linear trend, with Higher level students reporting a mean anxiety towards mathematics score that is almost one-fifth of a standard deviation below the OECD average; Ordinary level students have a mean around one-fifth of a standard deviation above the OECD average; and Foundation level students have a mean score that is one-half of a standard deviation above the OECD average.

Table 6.21. Mean Scores on Interest in Mathematics and Mathematics Anxiety, and Mean Score Differences, by Junior Certificate Mathematics Syllabus Level

	Interest in Mathematics				*Mathematics Anxiety*	
	%T	*%A*	*Mean*	*SE*	*Mean*	*SE*
Higher	42.6	42.9	0.19	0.03	-0.17	0.03
Ordinary	50.3	50.5	-0.23	0.03	0.21	0.02
Foundation	6.6	6.6	-0.26	0.06	0.52	0.08
Missing	0.5	0.0	0.31	0.23	0.09	0.14
All Available	99.5	100.0	-0.05	0.02	0.07	0.02

Mean Score Differences (Reference Category: Ordinary)

	Interest in Mathematics				*Mathematics Anxiety*			
	Diff	*SED*	*CI95L*	*CI95U*	*Diff*	*SED*	*CI95L*	*CI95U*
Found-Ord	0.0	0.06	-0.2	0.1	0.3	0.08	**0.1**	**0.5**
High-Ord	0.4	0.03	**0.3**	**0.5**	-0.4	0.04	**-0.5**	**-0.3**
Miss-Ord	0.5	0.23	0.0	1.1	-0.1	0.15	-0.5	0.2

Note. Total number of Irish students = 3880. N Higher = 1651, N Ordinary = 1941, N Foundation = 265, N Missing syllabus level = 24.
%T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

CONCLUSION

This chapter explored relationships between PISA 2003 and the Junior Certificate, in terms of performance on the two assessments in the three subject domains, and, in the case of mathematics only, the content of the two assessments.

A comparison of performance on Junior Certificate mathematics, English and science for males and females participating in PISA 2003 shows a pattern of gender differences consistent with that found for the population taking the Junior Certificate in 2003. However, the pattern of gender differences is not totally consistent when comparing PISA 2003 outcomes with those on the Junior Certificate. Males outperformed females on PISA 2003 mathematics, yet females do somewhat better than males on Junior Certificate mathematics. In contrast, there is no difference in the performance of males and females on PISA 2003 science, while the (small) gender difference associated with Junior Certificate science favours females. Comparing PISA 2003 reading with Junior Certificate English, the pattern is more consistent, in that females significantly outperform males on both assessments.

The overlap in performance on the three PISA 2003 assessment domains and the respective Junior Certificate subject areas is appreciable (with correlations between performance on each PISA domain and its corresponding Junior Certificate subject of around .70, and similar to 2000), but somewhat weaker than the overlap in performance on the PISA domains (where correlations among the three PISA domains range from .80 to .85). Clearly, many factors may be at play in mediating the relationship between performance on PISA and the Junior Certificate, not least differences in purposes, content and style of the assessments, and the manner in which results are scored and summarised.

The PISA 2003 mathematics assessment appears to be largely calculator neutral (which was explicitly intended according to the PISA 2003 assessment framework; OECD, 2003), with just over 70% of items rated by experts in Ireland as not requiring a calculator for their solution. That said, a significant association was found between calculator usage during the assessment and performance on PISA mathematics.

The chapter also explored the content and focus of PISA 2003 mathematics with respect to the Junior Certificate mathematics syllabus current at the time of the assessment. Ratings for PISA 2003 mathematics items in relation to the Junior Certificate syllabus suggest that three topic areas of the Junior Certificate (sets, geometry, and trigonometry) are not tapped at all by the PISA assessment. Junior Certificate algebra and relations and functions are also largely absent in PISA 2003 mathematics. PISA 2003 mathematics items that were rated as familiar fell most frequently into the Junior Certificate mathematics topic areas associated with applied arithmetic and measure, and statistics.

The distribution of curriculum familiarity ratings for the PISA 2003 mathematics items (which examined familiarity with concept, context of application, and item format for each syllabus level) are similar to those associated with the PISA 2000 assessment, despite the addition of two new mathematics areas (Quantity and Uncertainty were additions to the original two areas of Space & Shape and Change & Relationships). Students were expected to be familiar with concepts underlying half to seven-tenths of the PISA items depending on syllabus level. Familiarity with the context of application of these concepts and the item formats was lower, ranging from 65.9% to 80.0% unfamiliar with context and 62.4% to 83.5% unfamiliar with format. These ratings suggest that Foundation level students may find PISA mathematics to be particularly challenging.

The concept familiarity ratings were cross-tabulated with the four overarching ideas represented by the PISA mathematics subscales. Although one might have anticipated that differing patterns of familiarity across the four subscale areas (as rated by the Irish curriculum experts) might be associated with differences in student performance on those subscales, this is not the case. For example, students were moderately familiar with concepts underlying Space & Shape items, but performed significantly below the OECD average on this subscale. In contrast, students were, in general, least familiar with items associated with the Uncertainty subscale, yet performed significantly above the OECD average (see Chapter 3). Clearly, other factors may be at play, such as the interplay between concept familiarity and context familiarity, or the opportunities for students to learn concepts associated with PISA mathematics in contexts outside mathematics class.

The concept familiarity ratings were also cross-tabulated with the three PISA mathematics competency clusters. Students were expected, based on these ratings, to be most familiar with items associated with the Reproduction competency cluster (i.e., items involving the application of familiar procedures in routine mathematical settings) and least familiar with items associated with the Reflection competency cluster (i.e., items involving abstraction of mathematical information from non-mathematical contexts, mathematical argumentation and insight). This finding may be related to the observation that the Junior Certificate mathematics examination consists mainly of the application of learned procedures in abstract mathematical contexts, with little opportunity for the evaluation of students' application of mathematical skills in novel, real-world contexts.

Overall, the comparison of the Junior Certificate mathematics syllabus and PISA 2003 mathematics assessment shows marked differences in the manner in which such mathematical concepts are assessed. One might argue that Ireland's comparatively poor performance in mathematics relative to reading and science in PISA 2003 could be partly attributed to the fact that students were unfamiliar with many of the concepts tapped and also with the manner in which these concepts were contextualised (e.g., the setting of the problem in text and many of the item formats), as well as the types of mathematical reasoning involved (especially items associated with the Reflection cluster). Indeed, a comparison of some of the questions on the Junior Certificate mathematics examination papers from 2003 (Appendix A) with a selection of the PISA 2003 mathematics items (Appendix A) shows, at times, marked differences in the extent of abstraction, relational and conceptual understanding required for the PISA mathematics problems, relative to the requirements of the Junior Certificate mathematics examination. A particular difference is that, in PISA, students must extract relevant mathematical information from real-life contexts. This said, as there is no quantitative information on the nature and extent of match between PISA mathematics and the mathematics curricula in other countries, so findings relating to item familiarity cannot be interpreted from an internationally comparative perspective. Nonetheless, this preliminary analysis suggests that Irish students, while equipped with basic mathematical skills, may not be used to applying these skills in more challenging, non-routine contexts. If this pattern was found across a significant proportion of PISA 2003 mathematics items, it would suggest that there is general room for improvement in conceptual and relational understanding of core mathematics concepts. An in-depth analysis of the performance of Irish students on the PISA 2003 mathematics assessment at the item level is merited on this basis. The results from Chapter 3, which show that Ireland's average performance is characterised by comparatively high performance at the lower end of the achievement scale and comparatively low performance at the upper end, suggests that improvements in conceptual and relational understanding of higher achievers in particular could be made.

A comparison of the percentage of students taking each of the three Junior Certificate mathematics syllabus levels at each PISA mathematics proficiency level shows marked performance differences. Just over 70% of Foundation level students are at or below Level 1, and only 5.5% above Level 2. Even at Ordinary level, 36.2% of students are at Level 2, and over 21.9% at or below Level 1. At Higher level, almost 90% of students are at or above Level 2, but only 24.9% are at the highest levels (5 and 6). This stands in contrast to the overall performance of some countries in mathematics. For example, the percentage of *all* students participating in PISA in Belgium, Korea, Liechtenstein and the Netherlands scoring at Levels 5 or 6 equalled or exceeded 25%, the figure associated with Irish Higher level mathematics students. Comparatively low performance standards (as measured by PISA) at the upper end of the achievement distribution at Higher level, and more generally at Ordinary and Foundation levels, may be a cause for concern. The percentage of students scoring at or below Level 1 is higher than the percentages of students not achieving at least grade D on Junior Certificate mathematics at each syllabus level.

Similar comparisons were also made in the case of PISA reading literacy. At Higher level, almost half of students are at Levels 4 or 5, with just over 2% at or below Level 1; at Ordinary level, 27.1% are at or below Level 1, and at Foundation level, close to 80% are at or below Level 1. A comparison of the distribution of students taking each Junior Certificate English syllabus attaining each proficiency level in 2000 and 2003 indicates that at Higher level, fewer students in 2003 than in 2000 demonstrated the highest levels of proficiency. This is consistent with the comparatively low achievement of Irish students overall at the 75th, 90th, and 95th percentiles, noted in Chapter 3. In this respect, teachers of Junior Certificate English and their students might benefit from guidelines aimed at optimising the performance of higher achievers.

While the issue of familiarity with test content relates to the *fairness* of the assessment (that is, whether PISA 2003 mathematics assesses students' experiences of mathematics in second-level schools in Ireland), another relates to the *relevance* of both PISA 2003 mathematics and the Junior Certificate mathematics syllabus to young people. Issues of relevance raise questions such as: Which parts of these assessments might be considered more (and less) relevant to young people with various vocational and educational aspirations? Are there parts of the Junior Certificate mathematics syllabus which might be less relevant than others? Do all areas of the PISA 2003 mathematics assessment tap skills that mathematics educators and employers in Ireland consider relevant to young people? These in turn lead to further questions when one considers the performance on the PISA assessment as well as its content. For example, does the comparatively weak performance on the PISA Space & Shape subscale merit concern?

The large differences in the mean performance of students taking Junior Certificate mathematics at each syllabus level also raise concerns, even if they may in part reflect a mathematics syllabus that caters for wide differences in mathematics ability. In either case, the high proportions of Ordinary and Foundation students at or below Level 1 indicates that many of these students lack basic mathematics skills (or at least, the ability to implement such skills in the real life contexts used in PISA). The comparatively high anxiety towards mathematics reported by Ordinary and particularly Foundation level students, as well as the low interest in mathematics expressed by these students, should be taken into consideration and may merit more in-depth exploration.

A considered review of the content of the PISA mathematics assessment and the Junior Certificate mathematics examination is merited, and would be of particular value in the context of a broader review of mathematics education in Ireland. Of course, in addition to syllabus review, a related aspect, the teaching and learning of mathematics, may need to be considered. Any such review would need to take into account that the PISA 2003 assessment of mathematics is able to shed more light on performance in some Junior Certificate topic areas (applied arithmetic, measure and statistics), and considerably less in others (trigonometry, sets and geometry).

7

Achievement in Cross-Curricular Problem Solving

As indicated in the description of the assessment framework in Chapter 1, the problem-solving assessment was designed to assess students' ability to solve real-life problems set in contexts that differed from those used to assess mathematics, reading and science. In PISA, cross-curricular problem solving is defined as:

> . . . an individual's ability to use cognitive processes to confront and resolve real, cross-disciplinary situations where the solution path is not immediately obvious and where the literacy domains or curricular areas that might be applicable are not within a single domain of mathematics, science or reading. (OECD, 2003, p.156)

The assessment of cross-curricular problem solving incorporated three problem types: decision making (where students make decisions under specified constraints), system analysis and design (where students evaluate and design systems for a particular situation), and trouble shooting (where students identify why a device is malfunctioning, based on a set of symptoms). Nineteen problem-solving items set in units consisting of two to three items were presented. Problem-solving items appeared in 7 of the 13 PISA test booklets. Examples of items may be found in Appendix A, together with item-level scale values (at OECD Level), OECD average percent-correct scores, and percent-correct scores for Irish students. Where relevant, the percentages of students receiving partial and full credit on an item are given.

This chapter presents the results of the assessment of cross-curricular problem solving. First, performance of students in Ireland and other participating countries is described in terms of mean performance, performance at key markers such as the 10th and 90th percentiles, and performance across proficiency levels. Second, associations between performance in problem solving and performance in the other PISA domains are described. Third, between-school differences in problem solving are examined. Fourth, student-level variables associated with performance in problem solving, including student gender and socioeconomic status, are considered. Fifth, school-level variables associated with problem solving, including school type and school socioeconomic status, are examined. Sixth, correlations between continuous student- and school-level variables and problem solving are described.

OVERALL PERFORMANCE IN PROBLEM SOLVING

Mean Achievement

Performance on the assessment of cross-curricular problem solving was reported on a single scale with an OECD country average of 500 and a standard deviation of 100. As in the other minor domains, scores were imputed (inferred using statistical techniques) for students who were not asked to attempt the problem-solving items. Irish students ranked 18th of 29 OECD countries (95% confidence interval for Ireland's ranking = 17th to 19th), and 21st of 40 OECD and partner countries (95% confidence interval for ranking = 20th to 22nd), achieving a mean score (498.5) that was not significantly different from the OECD country average (Table 7.1; See Inset 3.1, and the description of Table 3.1 for a discussion of some of the issues that should be taken into account when interpreting tables such as Table 7.1). Sixteen countries achieved significantly higher mean scores than Ireland, including 13 OECD countries. Students in these countries had mean scores that ranged from 15 to 52 points higher than the mean score of Irish students. The standard deviation for Ireland (79.6) was the lowest among OECD countries, although the standard deviations for Finland (82.0), and Iceland (84.8) were close to Ireland's.

Table 7.1. Mean Achievement Scores and Standard Deviations on the Problem-Solving Scale – OECD and Partner Countries

	Pop	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*		*Pop*	*Mean*	*(SE)*	*SD*	*(SE)*	*OECD Diff*
Korea	⊠	550.4	(3.06)	86.4	(1.95)	▲	Luxembourg	⊠	493.66	(1.36)	91.6	(1.03)	▼
Hong Kong-Ch	⊠	547.9	(4.18)	97.2	(2.90)	▲	Slovak Rep	⊠	491.8	(3.38)	92.8	(2.39)	▼
Finland	⊠	547.6	(1.86)	82.0	(1.15)	▲	Norway	⊠	489.8	(2.60)	98.8	(1.65)	▼
Japan	⊠	547.3	(4.05)	104.9	(2.72)	▲	Poland	⊠	486.6	(2.78)	90.4	(1.68)	▼
New Zealand	⊠	532.8	(2.17)	95.7	(1.24)	▲	*Latvia*	⊠	482.5	(3.90)	92.1	(1.75)	▼
Macao-Ch	□	532.4	(2.53)	81.3	(2.55)	▲	Spain	⊠	482.2	(2.73)	93.6	(1.25)	▼
Australia	⊠	529.8	(1.98)	91.4	(1.35)	▲	*Russian Fed*	⊠	478.6	(4.59)	98.5	(2.11)	▼
Liechtenstein	□	529.5	(3.95)	92.7	(4.21)	▲	United States	⊠	477.3	(3.13)	98.1	(1.29)	▼
Canada	⊠	529.3	(1.74)	88.4	(0.93)	▲	Portugal	⊠	469.8	(3.87)	92.5	(2.11)	▼
Belgium	⊠	525.3	(2.20)	103.9	(1.52)	▲	Italy	⊠	469.5	(3.10)	102.2	(2.14)	▼
Switzerland	⊠	521.3	(3.05)	94.0	(1.88)	▲	Greece	⊠	448.5	(3.97)	98.8	(1.67)	▼
Netherlands	⊠	520.2	(2.95)	89.4	(2.03)	▲	*Thailand*	□	425.0	(2.72)	82.0	(1.59)	▼
France	⊠	519.2	(2.67)	92.9	(2.08)	▲	*Serbia and Monte.*	⊠	420.2	(3.32)	85.8	(1.56)	▼
Denmark	⊠	516.8	(2.54)	87.3	(1.51)	▲	*Uruguay*	×	410.7	(3.68)	111.7	(1.93)	▼
Czech Rep	⊠	516.4	(3.42)	92.9	(1.92)	▲	Turkey	×	407.5	(6.03)	96.7	(4.43)	▼
Germany	⊠	513.4	(3.24)	94.8	(1.75)	▲	Mexico	×	384.4	(4.30)	96.1	(2.01)	▼
Sweden	⊠	508.6	(2.44)	88.4	(1.58)	▲	*Brazil*	×	371.0	(4.84)	100.2	(2.60)	▼
Austria	⊠	506.1	(3.18)	90.0	(1.71)	o	*Indonesia*	×	361.4	(3.29)	73.3	(1.74)	▼
Iceland	⊠	504.7	(1.38)	84.8	(1.15)	▲	*Tunisia*	⊠	344.7	(2.11)	79.5	(1.42)	▼
Hungary	⊠	501.1	(2.86)	94.1	(2.03)	o	**OECD Total**		**489.7**	**(1.15)**	**106.4**	**(0.79)**	
Ireland	⊠	**498.5**	**(2.34)**	**79.6**	**(1.35)**	**o**	**OECD Average**		**500.0**	**(0.64)**	**100.0**	**(0.45)**	

⊠	>90% of 15-year olds enrolled	■	Mean significantly higher than Ireland	▲	Above OECD average
□	75-90% of 15-year olds enrolled	□	Mean not significantly different from Ireland	o	At OECD average
×	50-75% of 15-year olds enrolled	▒	Mean significantly lower than Ireland	▼	Below OECD average

Note. OECD countries are in regular font; partner countries are in italics. SD = Standard deviation; SE = Standard error.
The column "Pop" is an indicator of the percent of the 15-year-old population enrolled in schools in each country and is based on Column 15 of Table A3.1 in OECD (2004b).
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.

As with the other domains, nonparametric maximum likelihood estimation was used to classify countries into empirical groupings for problem solving (Table 7.2 and Figure 7.1; see Inset 3.2, for rationale for this method). Six groupings emerged. The top grouping consisted of four countries – Korea, Hong Kong-China, Finland and Japan. A second grouping consisted of 10 countries (one of them borderline.

Table 7.2. Six-point NPML Probability Distribution and Posterior Probabilities for the Problem-Solving Scale – OECD and Partner Countries

	Rank	Masspoints					
PISA score		*371.3*	*411.9*	*449.0*	*481.5*	*517.9*	*545.1*
Proportion		*0.136*	*0.164*	*0.039*	*0.269*	*0.257*	*0.135*
Korea	1	0.00	0.00	0.00	0.00	0.00	1.00
Hong Kong-China	2	0.00	0.00	0.00	0.00	0.00	1.00
Finland	3	0.00	0.00	0.00	0.00	0.01	0.99
Japan	4	0.00	0.00	0.00	0.00	0.01	0.99
New Zealand	5	0.00	0.00	0.00	0.00	0.54	0.46
Macao-China	6	0.00	0.00	0.00	0.00	0.58	0.42
Australia	7	0.00	0.00	0.00	0.00	0.78	0.22
Lichtenstein	8	0.00	0.00	0.00	0.00	0.80	0.20
Canada	9	0.00	0.00	0.00	0.00	0.81	0.19
Belgium	10	0.00	0.00	0.00	0.00	0.95	0.05
Switzerland	11	0.00	0.00	0.00	0.00	0.99	0.01
Netherlands	12	0.00	0.00	0.00	0.00	0.99	0.01
France	13	0.00	0.00	0.00	0.00	0.99	0.01
Denmark	14	0.00	0.00	0.00	0.00	1.00	0.00
Czech Republic	15	0.00	0.00	0.00	0.00	1.00	0.00
Germany	16	0.00	0.00	0.00	0.00	1.00	0.00
Sweden	17	0.00	0.00	0.00	0.01	0.99	0.00
Austria	18	0.00	0.00	0.00	0.04	0.96	0.00
Iceland	19	0.00	0.00	0.00	0.08	0.92	0.00
Hungary	20	0.00	0.00	0.00	0.35	0.65	0.00
Ireland	**21**	**0.00**	**0.00**	**0.00**	**0.65**	**0.35**	**0.00**
Luxembourg	22	0.00	0.00	0.00	0.95	0.05	0.00
Slovak Republic	23	0.00	0.00	0.00	0.98	0.02	0.00
Norway	24	0.00	0.00	0.00	0.99	0.01	0.00
Poland	25	0.00	0.00	0.00	1.00	0.00	0.00
Latvia	26	0.00	0.00	0.00	1.00	0.00	0.00
Spain	27	0.00	0.00	0.00	1.00	0.00	0.00
Russian Federation	28	0.00	0.00	0.00	1.00	0.00	0.00
United States	29	0.00	0.00	0.00	1.00	0.00	0.00
Portugal	30	0.00	0.00	0.02	0.98	0.00	0.00
Italy	31	0.00	0.00	0.02	0.98	0.00	0.00
Greece	32	0.00	0.00	0.99	0.00	0.00	0.00
Thailand	33	0.00	0.98	0.02	0.00	0.00	0.00
Serbia and Montenegro	34	0.00	1.00	0.00	0.00	0.00	0.00
Uruguay	35	0.00	1.00	0.00	0.00	0.00	0.00
Turkey	36	0.00	1.00	0.00	0.00	0.00	0.00
Mexico	37	0.98	0.02	0.00	0.00	0.00	0.00
Brazil	38	1.00	0.00	0.00	0.00	0.00	0.00
Indonesia	39	1.00	0.00	0.00	0.00	0.00	0.00
Tunisia	40	1.00	0.00	0.00	0.00	0.00	0.00

Probability of masspoint grouping is .95 or more; high degree of confidence

Probability of masspoint grouping is .90 to .95 or more; borderline degree of confidence

Probability of masspoint grouping is less than .90; low degree of confidence

OECD countries are in regular font; partner countries are in *italics* .

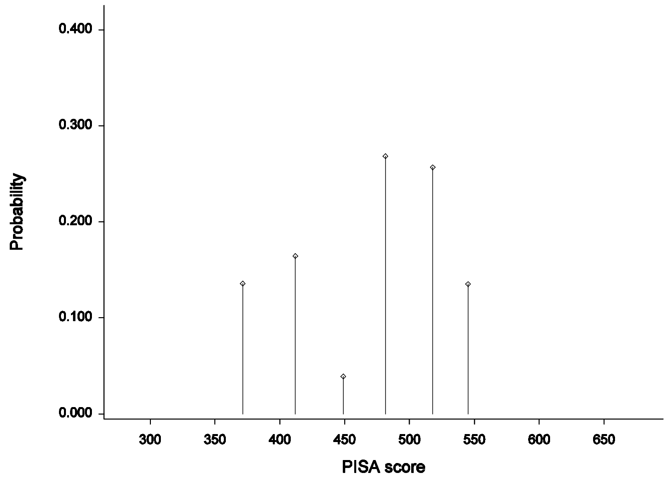


Figure 7.1. Plot of the NPML Probability Distribution for Problem-Solving Country Scores

Five countries has masspoints between the first and second groupings, and could not be reliably positioned in either. A third grouping consisted of a further 10 countries. Ireland had a masspoint between the second and third groupings, and could not be reliably assigned to one grouping or the other. The fourth grouping consisted of just one country. The fifth and sixth groupings each consisted of four countries. As in the other PISA 2003 domains, the country-level distribution of masspoints is negatively skewed.

Performance at Key Markers

Performance on PISA problem solving can be examined with reference to the scores achieved by students in Ireland and in other countries at key markers. Figure 7.2 shows the scores of Irish students scoring at the 10th, 25th, 75th, and 90th percentile ranks. For example, the score corresponding to the 10th percentile in Ireland is 394.6.

Scores for Ireland at the 75th, 90th, and 95th percentiles were lower, by approximately 15.5 points (one-sixth of a standard deviation), 24.3 points (one-quarter) and 31.1 points (three-tenths) respectively than the corresponding OECD country average scores (Table 7.3). On the other hand, Irish scores at the 5th, 10th, and 25th percentile ranks were higher by 37.5 points (one-third of a standard deviation), 26.1 points (one-quarter), and 10.8 points (one-tenth) respectively, than the OECD country average scores at the same markers.

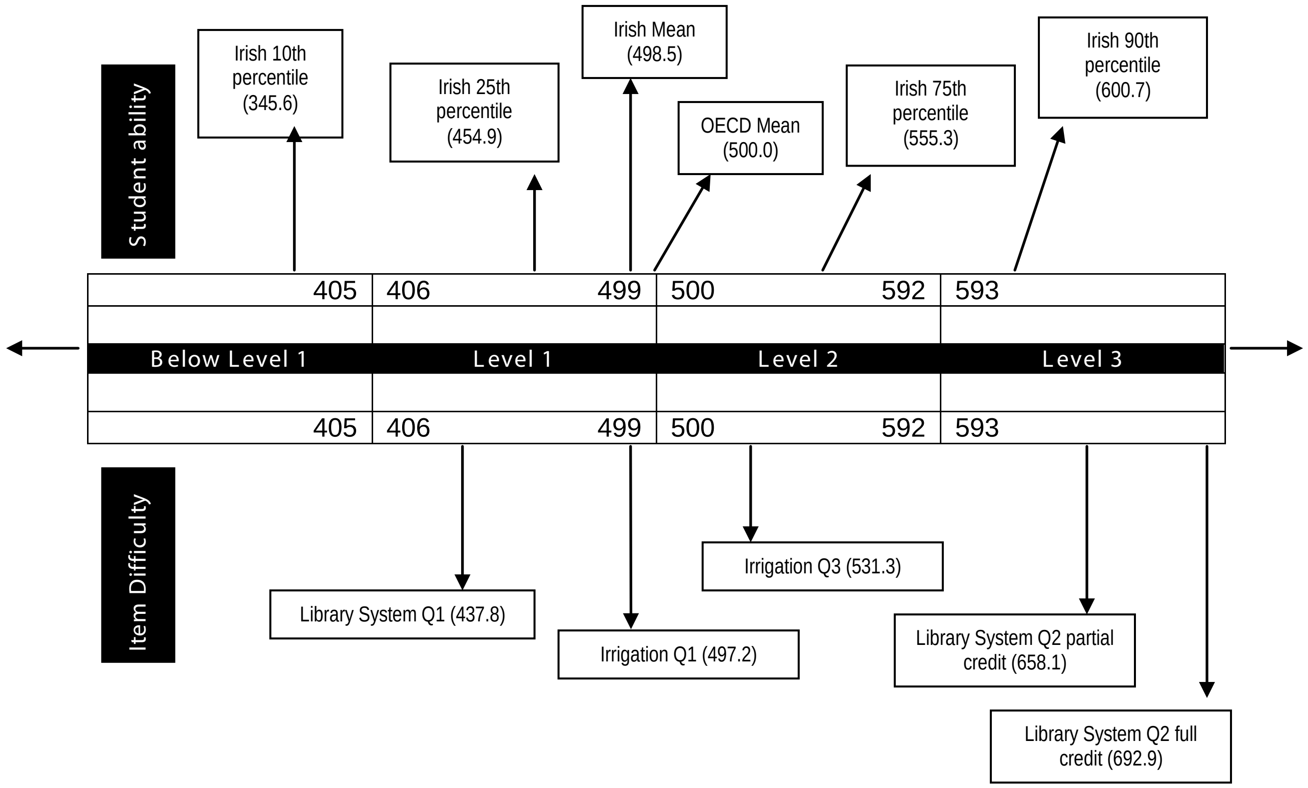
The highest-achieving Irish students also performed less well relative to their counterparts in countries with broadly similar mean achievement. For example, scores at the 90th and 95th percentiles in Norway and Sweden were higher, by about one-fifth of a standard deviation in each case, than the scores at these markers in Ireland (Table 7.3). On the other hand, whereas the scores for Sweden at the 5th and 10th percentile ranks were similar to those of Ireland, Norwegian scores at these markers were over one-third of a standard deviation lower.

Table 7.3. *Mean Scores of Students Achieving at the 5th, 10th, 25th, 75th, 90th and 95th Percentiles on the Problem-Solving Scale – Ireland and OECD Countries*

	5th		*10th*		*25th*		*75th*		*90th*		*95th*	
	Mean	*(SE)*	*Mean*	*(SE)*	*Mean*	*(SE)*	*Mean*	*(SE)*	*Mean*	*(SE)*	*Mean*	*(SE)*
Korea	404.1	(4.63)	437.8	(5.21)	493.9	(3.85)	609.9	(3.47)	658.3	(4.22)	686.0	(5.53)
HK-China	376.2	(10.53)	420.2	(7.86)	487.0	(6.12)	617.4	(3.18)	663.5	(2.93)	690.2	(3.72)
Finland	408.6	(4.67)	441.6	(2.80)	494.5	(2.48)	604.0	(2.28)	649.6	(2.33)	676.5	(3.58)
Japan	362.0	(8.27)	405.7	(6.83)	481.4	(5.71)	620.9	(4.23)	675.3	(4.56)	705.3	(6.01)
New Zealand	369.7	(3.75)	405.5	(4.23)	467.8	(3.65)	600.9	(2.38)	652.7	(2.49)	682.4	(2.76)
Macao-China	394.7	(6.41)	424.8	(5.58)	477.8	(3.72)	590.0	(4.33)	632.9	(5.43)	658.9	(6.54)
Australia	370.9	(4.07)	409.3	(3.46)	469.4	(2.78)	593.9	(2.13)	643.5	(2.67)	671.7	(3.41)
Liechtenstein	369.0	(14.93)	404.1	(11.09)	467.6	(6.03)	598.8	(9.30)	644.1	(10.46)	671.8	(12.04)
Canada	378.7	(2.43)	414.0	(2.78)	470.7	(2.46)	591.1	(1.89)	640.4	(2.13)	668.8	(2.40)
Belgium	340.3	(4.98)	382.5	(4.54)	455.5	(3.33)	601.8	(2.55)	653.3	(1.97)	681.0	(2.04)
Switzerland	358.1	(5.74)	397.0	(4.04)	461.3	(3.32)	586.9	(3.86)	636.9	(4.60)	666.2	(5.18)
Netherlands	372.1	(5.89)	401.5	(5.14)	456.1	(4.89)	586.6	(3.59)	636.4	(3.34)	662.1	(3.71)
France	358.3	(6.09)	395.8	(4.78)	458.6	(3.94)	585.8	(3.05)	634.7	(3.66)	662.3	(4.51)
Denmark	368.9	(5.02)	402.2	(4.28)	458.5	(3.06)	578.4	(2.76)	627.0	(3.44)	654.9	(3.71)
Czech Rep.	356.5	(8.56)	393.9	(6.17)	454.0	(4.43)	582.2	(3.58)	633.7	(3.87)	663.5	(4.01)
Germany	350.6	(5.91)	383.5	(5.35)	446.8	(4.78)	583.1	(4.28)	631.6	(2.66)	658.4	(3.19)
Sweden	359.8	(6.39)	395.1	(4.41)	451.3	(3.05)	570.8	(3.06)	619.0	(3.82)	647.2	(3.57)
Austria	356.8	(5.14)	387.9	(4.50)	443.0	(4.09)	569.4	(3.98)	620.7	(4.18)	650.9	(4.63)
Iceland	357.9	(5.47)	393.4	(3.28)	449.7	(2.22)	564.3	(2.03)	609.1	(2.33)	634.4	(3.59)
Hungary	342.9	(5.82)	377.6	(4.05)	436.4	(3.76)	566.7	(3.93)	622.1	(4.31)	653.1	(5.38)
Ireland	**364.0**	**(4.48)**	**394.6**	**(3.80)**	**444.9**	**(3.14)**	**555.3**	**(2.74)**	**600.7**	**(2.83)**	**624.5**	**(3.21)**
Luxembourg	339.1	(3.71)	373.4	(2.32)	432.0	(2.44)	558.0	(2.17)	609.8	(2.59)	639.9	(3.40)
Slovak Rep.	336.5	(7.13)	370.4	(5.86)	430.1	(4.71)	557.9	(3.58)	609.0	(3.84)	638.1	(4.16)
Norway	322.2	(5.54)	360.9	(4.64)	424.4	(3.72)	558.8	(3.28)	614.8	(4.17)	645.5	(4.38)
Poland	337.9	(5.55)	371.9	(4.07)	428.4	(3.10)	547.6	(2.97)	600.2	(3.53)	631.7	(4.49)
Latvia	326.2	(6.97)	361.9	(6.00)	419.8	(5.38)	547.2	(4.57)	599.0	(4.11)	628.4	(4.89)
Spain	321.8	(4.79)	361.2	(4.13)	420.9	(3.55)	547.3	(3.22)	599.4	(3.89)	629.4	(3.31)
Russian Fed.	314.2	(7.74)	350.9	(6.99)	412.6	(5.72)	546.0	(5.12)	603.7	(5.03)	637.2	(5.55)
USA	312.2	(5.62)	347.3	(4.55)	409.7	(4.07)	547.9	(3.30)	603.6	(3.96)	634.8	(4.19)
Portugal	311.2	(7.91)	345.2	(6.85)	408.6	(5.65)	534.3	(3.60)	586.3	(3.45)	614.5	(3.54)
Italy	288.7	(8.75)	334.1	(6.47)	406.1	(4.72)	540.1	(2.95)	595.4	(3.41)	627.1	(3.56)
Greece	283.1	(5.57)	319.2	(5.25)	382.7	(4.52)	516.5	(4.58)	574.5	(5.72)	607.5	(5.60)
Thailand	292.9	(3.87)	322.3	(3.37)	369.1	(2.64)	478.5	(3.99)	531.7	(4.01)	565.4	(5.97)
Serb. & Mont.	278.6	(4.15)	310.6	(4.40)	362.8	(3.94)	477.7	(4.19)	529.9	(4.93)	559.7	(5.08)
Uruguay	223.7	(5.69)	264.9	(5.10)	334.4	(4.66)	488.4	(5.45)	552.1	(5.04)	589.3	(5.74)
Turkey	257.1	(7.84)	290.6	(6.58)	343.0	(5.16)	466.5	(7.69)	531.1	(11.94)	576.7	(18.61)
Mexico	226.5	(5.44)	261.8	(5.18)	317.2	(5.20)	451.4	(5.07)	509.3	(5.69)	541.6	(6.48)
Brazil	211.0	(7.46)	244.1	(6.10)	301.8	(4.65)	437.8	(5.66)	500.6	(7.33)	538.4	(8.33)
Indonesia	244.6	(4.22)	269.8	(3.77)	311.7	(3.60)	408.6	(4.10)	456.8	(5.51)	486.9	(5.92)
Tunisia	212.9	(4.30)	242.6	(3.10)	290.6	(2.55)	399.9	(2.76)	445.8	(4.11)	474.4	(4.98)
OECD Total	**307.5**	**(2.73)**	**348.5**	**(2.16)**	**418.0**	**(1.73)**	**565.9**	**(1.29)**	**623.8**	**(1.33)**	**655.7**	**(1.45)**
OECD Avg.	**328.3**	**(1.65)**	**368.5**	**(1.31)**	**434.1**	**(1.08)**	**570.8**	**(0.85)**	**625.0**	**(0.79)**	**655.6**	**(0.80)**

Note. OECD countries are in regular font, partner countries are in *italics*. Countries are ordered in descending order of mean score on the problem solving scale.

Figure 7.2. The PISA Problem-Solving Scale: Cut-points for Proficiency Levels, Irish Scores at Key Markers, and Locations of Selected Items



Proficiency Levels

As in the domains of mathematics and reading, PISA established proficiency levels (intervals) for cross-curricular problem solving (see Inset 3.3, Chapter 3 for an explanation of how proficiency levels were developed). Four intervals were identified: Levels 3, 2, 1 and below Level 1. Students scoring at a particular level are expected to answer at least 50% of the items at that level correctly. Students at the bottom of a level have a .62 chance of correctly answering items at the bottom of the level (the easiest questions at that level), and a .42 chance of correctly answering items at the top of the level (the most difficult questions). Students at the top of a level have a .78 chance of answering the easiest items at that level correctly, and a .62 chance of answering the easiest items at the next level correctly. PISA does not describe the problem-solving skills of students with scores below Level 1. The top end of the proficiency scales is also unbounded (i.e., students scoring towards the top of Level 3 may have additional, higher-level problem-solving skills not assessed by PISA). Figure 7.2 shows the locations on the problem-solving scale of several of the sample problem-solving items presented in Appendix A as they relate to key markers such as the Irish and OECD mean scores.

Just 12.3% of Irish students achieved Level 3 on the proficiency scales in comparison with the OECD country average of 18.2% (Table 7.4). Students at this level are 'Reflective, communicative problem solvers', who can usually deal with a large number of constraints, and organise and monitor their thinking while developing solutions. The best students at this level can cope with multiple interrelated conditions that require students to work back and forth between their solution and the conditions laid out in the problem.

Just over 38% of Irish students achieved scores at Level 2, compared to the OECD country average of 34.2%. Students this level are described as 'reasoning decision-making problem solvers' who can typically apply analytic reasoning processes to solve problems requiring decision-making skills. Students at this level may need to combine various forms of representation (e.g., a formalised language, numerical information, and graphical information), handle unfamiliar representations (e.g. statements in a programming language or flow diagrams related to a mechanical or structural arrangement of components), or draw inferences based on two or more sources of information.

Table 7.4. Descriptions of Proficiency Levels on the Problem-Solving Scale, and Percentages of Students Achieving at Each Level – Ireland and OECD

		*Ireland**		*OECD***	
Level	*Brief Description*	*%*	*(SE)*	*%*	*(SE)*
Level 3 *(above 592.5)*	*"Reflective, communicative problem solvers"*: Approach multi-faceted problems systematically; construct their own representations; deal with a large number of interrelated conditions and constraints; organise and monitor their thinking while developing solutions; verify their solution; address problems successfully; and communicate their solutions clearly.	12.3	(0.76)	18.2	(0.18)
Level 2 *(498.9 to 592.5)*	*"Reasoning, decision-making problem solvers"*: Apply various types of reasoning (inductive deductive, and combinatorial) to analyse situations and solve problems among well-defined alternatives; combine and synthesise information and representations from a variety of sources; and draw inferences based on two or more sources of information.	38.3	(1.05)	34.2	(0.23)
Level 1 *(405.3 to 498.9)*	*"Basic problem solvers"*: Solve problems that deal with a single data source containing discrete, well-defined information; understand, locate and retrieve information on the major features of a problem; may transform information to present the problem differently, or apply information to check a limited number of well-defined conditions.	36.9	(1.17)	30.4	(0.20)
Below Level 1 *(less than 405.3)*	*"Weak or emergent problem solvers"*: Has a less than .50 chance of responding correctly to Level 1 tasks.	12.5	(0.91)	17.3	(0.25)
	Total	100.0		100.0	

* N (Ireland) = 3880. **Denotes OECD average percent.

Note. Students at a level have at least a 50% chance of correctly answering all items at that level.

Almost 37% of Irish students scored at Level 1 on the problem-solving scale compared to the OECD country average of 30.4%. Students scoring at this level can be categorised as 'Basic problem solvers' who can deal with a single data source containing well-defined information. They can usually transform information to represent the problem differently, and apply information to check a number of well-defined conditions. Just 12.5% of Irish students achieved scores below Level 1 compared to the OECD country average of 17.3%. Hence, in Ireland, there are comparatively fewer high and low achievers.

Table 7.5. *Percentage of Students at Each Proficiency Level on the Problem-Solving Scale – OECD and Partner Countries*

	< Level 1		*Level 1*		*Level 2*		*Level 3*	
	%	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*
Korea	5.2	(0.54)	21.6	(1.02)	40.8	(1.11)	32.4	(1.35)
Hong Kong-China	8.0	(1.05)	20.5	(0.99)	36.5	(1.17)	35.0	(1.39)
Finland	4.6	(0.51)	22.1	(0.84)	43.3	(0.82)	30.1	(0.92)
Japan	9.9	(1.00)	20.0	(0.98)	34.5	(1.25)	35.6	(1.58)
New Zealand	9.9	(0.81)	25.3	(0.85)	36.5	(1.01)	28.3	(0.92)
Macao-China	6.3	(0.82)	27.3	(1.37)	42.1	(2.00)	24.2	(1.59)
Australia	9.4	(0.59)	25.8	(0.67)	39.1	(0.83)	25.7	(0.84)
Liechtenstein	10.2	(1.54)	26.0	(2.35)	36.8	(3.56)	27.1	(2.62)
Canada	8.5	(0.48)	27.0	(0.70)	40.0	(0.69)	24.5	(0.72)
Belgium	13.6	(0.74)	24.4	(0.73)	33.7	(0.77)	28.3	(0.89)
Switzerland	11.4	(0.74)	26.8	(1.02)	38.7	(1.10)	23.1	(1.42)
Netherlands	10.7	(1.13)	30.5	(1.31)	35.8	(1.41)	23.0	(1.13)
France	11.7	(0.95)	28.1	(1.04)	37.5	(1.05)	22.7	(0.99)
Denmark	10.5	(0.76)	30.2	(0.92)	39.2	(0.90)	20.1	(0.93)
Czech Republic	12.1	(1.14)	29.4	(1.20)	37.0	(1.12)	21.5	(1.24)
Germany	14.2	(1.00)	27.7	(1.14)	36.4	(1.49)	21.7	(1.39)
Sweden	12.0	(0.85)	32.4	(1.12)	38.2	(1.02)	17.4	(0.99)
Austria	13.6	(1.00)	32.3	(1.07)	36.8	(1.06)	17.2	(1.16)
Iceland	12.4	(0.66)	32.5	(1.01)	40.2	(0.97)	14.9	(0.64)
Hungary	16.1	(0.98)	31.8	(1.40)	34.9	(1.20)	17.2	(1.16)
Ireland	**12.5**	**(0.91)**	**36.9**	**(1.17)**	**38.3**	**(1.05)**	**12.3**	**(0.76)**
Luxembourg	17.0	(0.71)	34.1	(0.95)	34.7	(1.04)	14.2	(0.56)
Slovak Republic	17.5	(1.40)	34.4	(1.16)	34.0	(1.34)	14.1	(0.96)
Norway	19.4	(0.88)	32.6	(1.15)	33.1	(0.97)	14.9	(0.84)
Poland	17.5	(0.98)	37.2	(1.02)	33.6	(1.14)	11.7	(0.66)
Latvia	20.3	(1.48)	35.6	(1.27)	32.5	(1.39)	11.6	(0.96)
Spain	20.1	(0.87)	35.5	(1.08)	32.9	(1.15)	11.6	(0.83)
Russian Federation	22.8	(1.69)	34.5	(1.03)	30.6	(1.28)	12.2	(1.02)
United States	23.7	(1.12)	33.7	(0.83)	30.3	(1.03)	12.3	(0.78)
Portugal	23.9	(1.73)	36.5	(1.09)	31.0	(1.36)	8.6	(0.63)
Italy	24.7	(1.32)	34.7	(1.16)	30.0	(0.99)	10.6	(0.66)
Greece	32.7	(1.53)	36.1	(1.01)	24.3	(1.21)	7.0	(0.80)
Thailand	41.4	(1.57)	40.5	(1.08)	15.6	(1.10)	2.6	(0.47)
Serbia and Montenegro	42.6	(1.70)	39.5	(1.16)	15.8	(1.17)	2.1	(0.31)
Uruguay	47.2	(1.58)	30.5	(1.33)	17.5	(1.23)	4.7	(0.54)
Turkey	51.2	(2.51)	32.5	(1.58)	12.4	(1.64)	3.9	(1.24)
Mexico	58.1	(1.86)	29.7	(1.07)	10.9	(0.99)	1.3	(0.24)
Brazil	64.1	(1.92)	25.6	(1.52)	8.7	(1.12)	1.6	(0.48)
Indonesia	73.5	(1.71)	22.9	(1.36)	3.5	(0.64)	0.1	(0.07)
Tunisia	77.1	(1.05)	20.4	(0.82)	2.5	(0.51)	0.1	(0.05)
OECD Total	**21.6**	**(0.44)**	**30.0**	**(0.28)**	**31.2**	**(0.41)**	**17.2**	**(0.33)**
OECD Average	**17.3**	**(0.25)**	**30.4**	**(0.20)**	**34.2**	**(0.23)**	**18.2**	**(0.18)**

Note. OECD countries are in regular font, partner countries are in *italics* . Countries are ordered in descending order of mean score on problem solivng.

ASSOCIATIONS BETWEEN PERFORMANCE IN PROBLEM SOLVING AND IN OTHER PISA DOMAINS

As indicated in Chapter 3, correlations ranging from .80 to .85 were observed between performances on the PISA domains of mathematics, reading and science (Irish student achievement). In this section, associations between performance on these domains and on cross-curricular problem solving are examined. The correlation coefficients considered here have not been adjusted for attenuation.

Of particular interest is the correlation between mathematics and problem solving. Although it is acknowledged that there is overlap between the two domains in terms of the problem-solving processes in which students are expected to engage, it has been argued that the domains are independent of one another to the extent that the problem-solving assessment makes minimum demands on students' mathematical knowledge (OECD, 2004c). It is noteworthy, therefore, that the correlation between mathematics and problem solving for Irish student achievement is .90 (Table 7.6).

It is also noteworthy that there are strong correlations between reading and problem solving (.87), and between science and problem solving (.85). Although the PISA problem-solving items might have been designed to be different from those encountered in traditional school domains, problem solving shares considerable variance with the other PISA domains.

Table 7.6. *Linear Associations Between Combined Mathematics, Reading, Science and Problem-Solving Scales (Irish Students)*

	Raw Coeff	*SD Unit*	*SE*	*r*	*t*	*p*
Maths-Prob. Solving	0.84	66.80	0.01	**.899**	74.61	<.001
Reading-Prob. Solving	0.80	63.38	0.01	**.866**	82.70	<.001
Science-Prob. Solving	0.72	57.53	0.01	**.845**	143.77	<.001

Note. The column "SD Unit." shows the increase in the outcome variable per standard deviation increase in the explanatory variable.
Standard deviations for scales are as follows: Combined mathematics - 85.3, Reading - 87.5, Science - 93.0, Problem Solving - 79.6.

BETWEEN-SCHOOL DIFFERENCES IN PROBLEM-SOLVING OUTCOMES

A comparison of the proportion of total variation between schools in achievement in mathematics, reading, and science suggested that the Irish system is comparatively homogenous; i.e., compared to other countries, schools in Ireland do not differ greatly with respect to achievement (see Chapter 5). The same is true of the between-school differences in problem solving, with just 15.7% of the total variation lying between schools in Ireland, compared to an OECD average of 31.6% (country range = 3.0% to 57.6%).

PROBLEM SOLVING AND STUDENT CHARACTERISTICS

In this section, mean differences in problem solving associated with selected student variables are considered. The variables are: gender, socioeconomic status, absence from school, lone-parent status, number of siblings, home educational resources, books in the home, current grade level, study of science, self-efficacy in mathematics, and anxiety about mathematics. Readers can compare results here with results for the other PISA domains described in Chapter 4, where information is also provided on how explanatory variables were constructed, and how the tables presented here can be interpreted (Insets 4.1, 4.2 and 4.3; see also Appendix B, Section B.1).

Gender

Male students performed marginally better than female students on problem solving (Table 7.7). The mean score difference (one-half of one point) is not statistically significant. This contrasts with the finding of a significant difference of one-sixth of a standard deviation in favour of male students in mathematics (see Table 4.1, Chapter 4).

Table 7.7. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Gender

	%T	*Mean*	*SE*
Female	49.7	498.2	3.46
Male	50.3	498.7	2.78
All Available	100.0	500.5	2.34

Mean Score Differences (Reference Category: Male)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
Female–Male	-0.5	4.20	-8.9	7.8

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals.

The OECD average difference between males and females on problem solving was not statistically significant. However, female students significantly outperformed male students in four OECD countries – Iceland, Finland, Sweden and Norway – and in two partner countries – Thailand and Indonesia. In contrast, male students outperformed females in partner country Macao-China only. The largest difference – in favour of females in Iceland – was 30.5 points (OECD, 2004c, Table 5.1).

A comparison of the performance of Irish male and female students at key markers (5th, 10th, 25th, 75th, 90th and 95th percentile ranks) revealed marginally higher scores for males at the three lowest percentile ranks, and marginally lower scores at the three highest percentile ranks. None of the differences, however, is statistically significant (Table 7.8). The percentages of males performing below Level 1, and at Level 3, on the problem-solving scale, are marginally greater than the percentage of females, while the opposite pattern is found for Levels 1 and 2 (Table 7.9). Again, none of these differences is significant.

Table 7.8. Mean Problem-Solving Scores of Irish Students at Six Key Markers, and Mean Score Differences, by Gender

	Males		*Females*		*All*	
Level	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
5th	366.4	5.31	361.6	6.14	363.8	4.30
10th	396.0	4.73	392.5	5.12	394.5	3.65
25th	444.9	3.86	444.8	3.95	444.8	3.04
75th	554.3	4.30	555.8	3.30	555.2	2.69
90th	598.3	4.42	602.3	3.67	600.4	2.61
95th	620.7	4.50	626.7	4.45	624.4	3.04

Mean Score Difference (Reference Category: Male)

	Difference	*SED*	*CI95L*	*CI95U*
F5th-M5th	4.81	8.12	-16.6	26.2
F10th-M10th	3.57	6.97	-14.8	22.0
F25th-M25th	0.06	5.52	-14.5	14.6
F75th-M75th	-1.58	5.42	-15.9	12.7
F90th-M90th	-3.98	5.75	-19.1	11.2
F95th-M95th	-6.01	6.32	-22.7	10.7

Note. SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted confidence intervals

Table 7.9. Percentages of Irish Students at Each Problem-Solving Proficiency Level, and Percentage Differences, by Gender

	Males		*Females*		*All*	
Level	*Percent*	*SE*	*Percent*	*SE*	*Percent*	*SE*
Below Level 1	12.8	1.55	12.2	1.55	12.5	0.91
Level 1	36.6	2.12	37.2	3.13	36.9	1.17
Level 2	37.6	2.16	39.0	2.02	38.3	1.05
Level 3	13.0	0.77	11.6	1.72	12.3	0.76
Total	100.0		100.0		100.0	

Percentage Difference (Reference Category: Male)

	Difference	*SED*	*CI95L*	*CI95U*
Below Level 1	-0.6	2.19	-6.0	4.9
Level 1	0.6	3.78	-8.8	10.0
Level 2	1.4	2.96	-6.0	8.8
Level 3	-1.4	1.88	-6.1	3.3

Note. SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted confidence intervals

Socioeconomic Status

Using the measure of socioeconomic status (SES) described in Chapter 4, based on the highest of either parent's main occupation, the mean scores in problem solving of students categorised as high, medium, and low in socioeconomic status were compared. Students with high SES achieved a mean score that was 29.0 points (one-third of a standard deviation) higher than the mean score of students with medium SES, and 61.3 points (four-fifths of a standard deviation) higher than the mean score of students with low SES (Table 7.10). The pattern is broadly similar to the pattern of findings for mathematics, reading, and science (see Table 4.8).

Table 7.10. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Parental Occupation (SES)

	%T	*%A*	*Mean*	*SE*
Low	31.0	32.4	469.3	3.30
Medium	33.6	35.2	501.6	2.56
High	31.1	32.5	530.6	3.25
Missing	4.3	0.0	452.7	15.83
All Available	95.7	100.0	500.5	2.25

Mean Score Differences (Reference Category: High)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
Low-High	-61.3	4.64	**-72.7**	**-50.0**
Medium-High	-29.0	3.98	**-38.7**	**-19.3**
Missing-High	-77.8	16.23	**-117.5**	**-38.1**

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals.

Family Structure – Lone-Parent Status

Students living in lone-parent households achieved a significantly lower mean score in problem solving than students living in dual-parent households (i.e., nuclear/mixed families) (Table 7.11). The difference – 29.7 points (two-fifths of a standard deviation) – was similar in size to the differences for mathematics, reading, and science, where students in lone-parent families also performed significantly less well (Table 4.11).

Table 7.11. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Lone-Parent Status

	%T	*%A*	*Mean*	*SE*
Lone parent	15.1	15.7	474.0	4.07
Dual parent	81.3	84.3	503.7	2.40
Missing	3.6	0.0	483.8	14.21
All Available	96.4	100.0	499.0	2.37

Mean Score Differences (Reference Category: Lone Parent)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
Dual–Lone	29.7	3.92	**20.8**	**38.6**
Missing–Lone	9.9	15.18	-24.8	44.6

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Family Structure – Number of Siblings

The mean achievement scores in problem solving of students with varying numbers of siblings were compared (Table 7.12). Students with one sibling outperformed students with no siblings (by 20.8 points or over one-quarter of a standard deviation), two siblings (by 10.3 points or one-eighth of a standard deviation), three siblings (by 24.1 points or three-tenths of a standard deviation), and four or more (by 39.9 points or one-half of a standard deviation). Results are broadly similar to those for mathematics, reading and science. However, problem solving was the only domain in which students with one sibling had a statistically significantly higher mean score than students with two siblings.

Table 7.12. Mean Problem-Solving Scores of Irish Students and Mean Score Differences, by Number of Siblings

	%T	*%A*	*Mean*	*SE*
0 Sibling	10.6	10.8	497.5	4.41
1 Sibling	16.9	17.3	518.3	3.79
2 Siblings	25.6	26.1	507.9	3.21
3 Siblings	20.7	21.2	494.1	3.47
≥4 Siblings	24.0	24.5	478.4	3.54
Missing	2.2	0.0	500.1	21.17
All Available	97.8	100.0	498.4	2.34

Mean Score Differences (Reference Category: One)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
None–one	-20.8	5.13	**-34.3**	**-7.3**
Two–one	-10.3	3.70	**-20.1**	**-0.6**
Three–one	-24.1	4.49	**-36.0**	**-12.3**
≥Four–one	-39.9	4.79	**-52.5**	**-27.2**
Missing–one	-18.2	21.58	-75.2	38.7

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Home Educational Resources

As indicated in Chapter 4, students were categorised with regard to the number of educational resources in the home. Students with a desk for study, a quiet place to study, and books to help with schoolwork were categorised as having a high level of educational resources. Those with two of these were categorised as medium, while those with one or none were categorised as low. Students with high educational resources achieved a mean score in problem solving that was 19.0 points (just under one-quarter of a standard deviation) higher than the mean score of students with medium resources, and 42.5 points (over one-half of a standard deviation) higher than the mean score of students with low resources (Table 7.13).

Table 7.13. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Educational Resources in the Home

	%T	*%A*	*Mean*	*SE*
Low	17.9	18.1	468.1	4.36
Medium	24.7	24.9	491.5	2.92
High	56.4	57.0	510.6	2.77
Missing	1.0	0.0	531.4	23.45
All Available	99.0	100.0	498.1	2.34

Mean Score Differences (Reference Category: High)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
Low–High	-42.5	4.88	**-54.4**	**-30.6**
Medium–High	-19.0	3.31	**-27.1**	**-10.9**
Missing–High	20.8	23.91	-37.6	79.3

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Books in the Home

As in the other PISA assessment domains, students with more books in the home outperformed students with fewer books on the assessment of problem solving (Table 7.14; see also Table 4.14, Chapter 4). Students with 26 to 100 books achieved a mean score that was significantly higher, by 51.8 points, than the mean score of students with zero to 10 books, and significantly lower than the mean scores of students with at least 100 books. The difference in favour of students with more than 500 books over those with 26 to 100 was 45.9 points (more than one half of a standard deviation).

Absence from School

Students who reported that they had not missed school in the two weeks prior to taking the PISA assessment achieved a mean score on problem solving that was significantly higher, by 44.4 points (just over one-half of a standard deviation) than the mean score of students who were absent on three or more days (Table 7.15). Students with no absences also had a mean score that was significantly higher, by 16.7 points (about one-fifth of a standard deviation), than the mean score of students who were absent on one or two days. Again, these findings are similar to the findings for the other assessment domains (see Table 4.20).

Table 7.14. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Index of Books in the Home

	%T	*%A*	*Mean*	*SE*
0 to 10 books	10.3	10.5	444.9	5.33
11 to 25 books	14.8	15.1	465.5	3.80
26 to 100 books	32.8	33.5	496.6	2.74
101 to 200 books	19.5	19.9	522.5	2.88
201 to 500 books	13.4	13.6	524.3	4.13
>500 books	7.3	7.4	542.5	4.75
Missing	2.1	0.0	481.2	23.38
All Available	97.9	100.0	498.8	2.32

Mean Score Differences (Reference Category: 26-100)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
0 to 10–26 to 100	-51.8	5.56	**-66.8**	**-36.7**
11 to 25–26 to 100	-31.1	3.90	**-41.6**	**-20.5**
101 to 200–26 to 100	25.9	3.60	**16.1**	**35.6**
201 to 500–26 to 100	27.7	4.58	**15.3**	**40.1**
>500–26 to 100	45.9	5.21	**31.8**	**60.0**
Missing–26 to 100	-15.4	23.54	-79.1	48.3

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Table 7.15. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Level of Absenteeism

	%T	*%A*	*Mean*	*SE*
None	58.2	60.2	508.9	2.48
1 or 2	28.8	29.8	492.1	3.03
≥3	9.7	10.0	464.5	5.81
Missing	3.3	0.0	469.9	17.13
All Available	96.7	100.0	499.4	2.28

Mean Score Differences (Reference Category: None)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
1 or 2–None	-16.7	2.95	**-23.9**	**-9.5**
≥3 – None	-44.4	5.84	**-58.7**	**-30.1**
Missing–None	-39.0	16.88	-80.3	2.3

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Current Grade Level

Irish students in Third year achieved a mean score that was significantly higher, by nine-tenths of a standard deviation (72.3 points), than the mean score of students in Second year, and significantly lower, by two-thirds of a standard deviation (54.9 points), than the mean score of students in Fourth (Transition) year (Table 7.16). As in the other PISA domains (see Table 4.19), the gap in problem-solving performance between students in Third and Fifth year is smaller than the gap between students in Third and Fourth year.

Table 7.16. Mean Problem-Solving Scores of Irish Students and Mean Score Differences, by Current Grade (Year) Level

	%T	*Mean*	*SE*
2nd Year	2.8	414.8	8.47
3rd Year	60.9	487.2	2.79
4th Year	16.7	542.0	4.08
5th Year	19.6	508.4	4.25
All Available	100.0	498.5	2.34

Mean Score Differences (Reference Category: Third Year)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
2nd–3rd	-72.3	8.87	**-94.0**	**-50.7**
4th–3rd	54.9	4.47	**43.9**	**65.8**
5th–3rd	21.2	4.68	**9.7**	**32.6**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

Study of Science for the Junior Certificate

In Chapter 4, it was observed that students who reported that they did not study science for the Junior Certificate Examination achieved mean scores in mathematics, reading and science that were significantly lower than the mean scores of students who studied science (Table 4.24). Students who indicated that they did not study science (9.9% of the sample) also performed significantly less well on problem solving than did students who studied the subject. The difference between the groups was 48.7 points (over three-fifths of a standard deviation) (Table 7.17).

Table 7.17. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Study of Science at Junior Certificate Level

	%T	*%A*	*Mean*	*SE*
Yes	88.4	89.9	505.2	2.17
No	9.9	10.1	456.5	5.49
Missing	1.7	0.0	391.7	12.33
All Available	98.3	100.0	500.3	2.26

Mean Score Differences (Reference Category: Yes)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
No−Yes	-48.7	5.55	**-61.3**	**-36.0**
Missing−Yes	-113.5	12.48	**-142.1**	**-85.0**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Self-Efficacy in Mathematics

In Chapter 4, students' self-efficacy in mathematics was found to be associated with their performance in the domain of mathematics. Students who indicated high levels of confidence in their ability to solve a range of mathematics problems were found to perform significantly better than students reporting medium and low levels of confidence. The mean score difference in mathematics between students in the high and low categories was about one and one-quarter standard deviations (see Table 4.25).

When students' scores on the assessment of cross-curricular problem solving were related to their self-efficacy in mathematics, students with high levels of self-efficacy were observed to have a mean score that was 47.2 points (almost three-fifths of a standard deviation) higher than the mean score of students with medium levels, and 91.7 points (one and one-sixth standard deviations) higher than the mean score of students with low levels (Table 7.18).

Table 7.18. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Perceived Self-Efficacy in Mathematics

	%T	*%A*	*Mean*	*SE*
Low	30.4	30.9	454.2	2.65
Medium	38.9	39.6	498.8	2.73
High	29.0	29.5	546.0	2.63
Missing	1.6	0.0	471.0	34.18
All Available	98.4	100.0	498.9	2.30

Mean Score Differences (Reference Category: High)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
Low-High	-91.7	3.19	**-99.5**	**-83.9**
Medium-High	-47.2	3.00	**-54.5**	**-39.9**
Missing-High	-74.9	34.19	-158.5	8.7

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals.

Anxiety About Mathematics

As indicated in Chapter 4, students reporting high levels of anxiety about mathematics did significantly less well on the combined mathematics scale than students with medium and low levels (Table 4.26). Students with high levels of anxiety about mathematics also did less well on problem solving (Table 7.19). The mean score difference between students with low levels of anxiety over those with high levels is 58.3 points (over seven-tenths of a standard deviation). This is similar to the mean score difference in mathematics.

Table 7.19. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Anxiety About Mathematics

	%T	*%A*	*Mean*	*SE*
Low	30.7	31.3	527.5	3.10
Medium	39.8	40.5	498.1	2.91
High	27.7	28.2	469.2	3.00
Missing	1.8	0.0	462.2	33.22
All Available	98.2	100.0	499.1	2.28

Mean Score Differences (Reference Category: High)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
Low-High	-58.3	4.10	**-68.3**	**-48.3**
Medium-High	-29.4	3.55	**-38.1**	**-20.7**
Missing-High	-65.3	33.10	-146.2	15.6

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

PROBLEM SOLVING AND SCHOOL CHARACTERISTICS

In this section, performance on PISA problem solving is related to selected school-level variables. The variables are school sector, Junior Certificate Examination fee waiver, and disciplinary climate.

School Sector

In problem solving, students attending secondary schools significantly outperformed students in community/comprehensive schools by 16.4 points (just over one-fifth of a standard deviation), and students in vocational schools by 38.1 points (almost half a standard deviation) (Table 7.20). These are similar to the achievement differences associated with the other PISA assessment domains (Table 4.28, Chapter 4).

Table 7.20. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by School Sector

	%T	*Mean*	*SE*
Community/Comprehensive	17.3	493.2	4.50
Secondary	61.0	509.6	3.06
Vocational	21.7	471.5	5.65
All Available	100.0	498.5	2.34

Mean Score Differences (Reference Category: Secondary)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
Community–Secondary	-16.4	5.45	**-28.9**	**-3.9**
Vocational–Secondary	-38.1	6.44	**-52.8**	**-23.4**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Junior Certificate Examination Fee Waiver Entitlement

The percentage of 15-year-olds in a school entitled to a fee waiver in the Junior Certificate Examination was used as a measure of school-level SES. In problem solving, students attending schools with low fee-waiver entitlement achieved a mean score that was significantly higher than the mean score of students in schools with medium entitlement by 23.8 points (three-tenths of a standard deviation), and the mean score of students in schools with a high entitlement by 55.3 points (about seven-tenths of a standard deviation) (Table 7.21).

Table 7.21. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Junior Certificate Examination Fee Waiver Entitlement

	%T	*Mean*	*SE*
Low	32.8	524.8	4.06
Medium	34.2	501.0	3.11
High	33.0	469.5	4.45
All Available	100.0	498.5	2.34

Mean Score Differences (Reference Category: Low)

	Problem Solving			
	Diff	*SED*	*CI95L*	*CI95U*
Medium- low	-23.8	5.35	**-36.0**	**-11.6**
High- low	-55.3	6.25	**-69.6**	**-41.0**

Note. N = 3880. %T = percentage all; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences (p = .05) are highlighted in bold.

Disciplinary Climate

Students attending schools in which they perceived the disciplinary climate in mathematics classes to be high (positive) achieved a mean score in problem solving that was significantly higher, by 19.6 points (over one-quarter of a standard deviation), than the mean score of students attending schools with medium levels, and significantly higher, by 37.6 points (about one half of a standard deviation), than the mean score of students attending schools with low (negative) levels (Table 7.22).

Table 7.22. Mean Problem-Solving Scores of Irish Students, and Mean Score Differences, by Disciplinary Climate in Mathematics Class

	%T	*%A*	*Mean*	*SE*
Low	29.2	29.8	480.8	3.66
Medium	40.4	41.3	499.9	2.57
High	28.3	28.9	518.5	3.49
Missing	2.1	0.0	447.7	30.80
All Available	97.9	100.0	499.6	2.25

Mean Score Differences (Reference Category: High)

	Diff	*SED*	*CI95L*	*CI95U*
Low–High	-37.6	4.77	**-49.3**	**-26.0**
Medium–High	-18.6	3.24	**-26.5**	**-10.7**
Missing–High	-70.8	31.24	-147.1	5.6

Note. N = 3880. %T = percentage all; %A = percentage available; Diff = mean difference; SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted 95% confidence intervals. Confidence intervals for significant differences ($p \leq .05$) are highlighted in bold.

CORRELATIONS BETWEEN EXPLANATORY VARIABLES AND PERFORMANCE ON PROBLEM SOLVING

In this section, linear associations (correlations) between continuous explanatory variables at the student and school levels and performance on PISA problem solving are presented. Inset 4.4 provides information on the interpretation of correlation coefficients.

Student-Level Variables

The student-level variable with the strongest relationship to performance in problem solving was self-efficacy in mathematics (r = .47) (Table 7.23). Both the index of the number of books in the home (.35) and socioeconomic status (.34) correlate moderately with problem solving. The correlation between home educational resources and problem solving (.22) is somewhat weaker.

The correlation between anxiety about mathematics and problem solving (–.33) is significant and negative, indicating that students with higher levels of anxiety performed less well on problem solving than students with low levels. A similar outcome was observed for mathematics (Table 4.36). Absence from school is also negatively correlated with achievement (–.18), indicating that students with more absences tended to perform less well than students with fewer absences.

Table 7.23. Linear Associations between Student Variables and Problem Solving

	r	*t*	*p*
Number of siblings	**-.120**	-5.85	<.001
Socioeconomic status (parent ed)	**.336**	14.28	<.001
Home educational resources	**.223**	10.19	<.001
Number of books in the home	**.348**	16.65	<.001
Absence from school	**-.175**	-7.91	<.001
Self-efficacy in mathematics	**.469**	29.73	<.001
Anxiety about mathematics	**-.325**	-17.62	<.001

Note. Significant correlations are highlighted in bold.
Df = 80 (number of variance strata associated with balanced repeated replicate (BRR) method of variance estimation).

School-Level Variables

Just two continuous variables were selected for analysis in relation to problem solving – percentage of Junior Certificate Examination students in the school with fee-waiver entitlement, and disciplinary climate in mathematics classes (Table 7.24). The correlation between fee-waiver entitlement and problem solving (.31) is similar to those reported for mathematics, reading and science (Table 4.38). The weak to moderate correlation between disciplinary climate in mathematics classes and problem solving (.19) is close to that reported for disciplinary climate and performance on the other assessment domains.

Table 7.24. Linear Associations between School Variables and Problem Solving

	r	*t*	*p*
Fee waiver entitlement	**.306**	-9.69	<.001
Disciplinary climate in maths lessons	**.190**	8.17	<.001

Note. Significant correlations are highlighted in bold.
Df = 80 (number of variance strata associated with balanced repeated replicate (BRR) method of variance estimation).

CONCLUSION

This chapter reported the outcomes of the assessment of cross-curricular problem solving in PISA 2003. Problem solving was assessed using 19 problems situated in real-life contexts that called on students to apply analogical reasoning skills. Three problem types – decision making, systems analysis and design, and trouble shooting – were included.

Irish students achieved a mean score (498.5) on problem solving that did not differ significantly from the OECD country average of 500. Ireland ranked 18th of 29 OECD countries (95% confidence interval of Ireland's ranking = 17th to 19th), and 21st of 40 OECD and partner countries (95% confidence interval of Ireland's ranking = 20th to 22nd). This similar to that for mathematics (20th). Some countries showed greater variation in their rankings across the two domains. Whereas the Netherlands ranked fourth in mathematics, its rank was twelfth in problem solving. New Zealand, on the other hand, ranked twelfth in mathematics and fifth in problem solving.

When nonparametric maximum likelihood estimation was employed to group countries on the basis of their overall performance on problem solving, Ireland (and Hungary) had masspoints between the second and third groupings of countries.

Irish students at the 75th, 90th and 95th percentile ranks achieved scores below than the corresponding OECD country average scores. In contrast, Irish students scoring at the 5th, 10th, and 25th percentile ranks achieved scores that were higher. A similar finding emerged when performance across the problem-solving proficiency levels was examined. Just 12.3% of Irish students achieved Level 3 (the highest level) in comparison with an OECD country average of 18.2%. In contrast, 12.5% of Irish students achieved scores below Level 1 (the lowest level), compared to the OECD country average of 17.3%.

The performance of Irish students in problem solving was remarkably similar to their performance on the combined mathematics scale. In both domains, they achieved mean scores that are not significantly different from the corresponding OECD country averages, and differences between high and low achievers are comparatively small. Moreover, there is a strong correlation (.90) between mathematics and problem solving, even though the problem-solving items had been designed so that students would only require minimal mathematics knowledge. Correlations between problem solving and the other PISA domains are marginally weaker (.85 for both reading and science).

An analysis of the association between a range of student- and school-level variables and performance on problem solving produced findings similar to those reported for mathematics in Chapter 4 with a the exception those relating to gender. Unlike the domains of mathematics, reading and science, however, school- and student-level variables have not been examined simultaneously in a multilevel model of achievement. Such models can reveal somewhat different patterns of association than those reported in this chapter, as the effects of a particular variable can be estimated while adjusting for the effects of others in the model.

The differences in performance between males and females in problem solving for overall mean scores, and scores at key percentile points and proficiency levels, are not statistically significant. Differences were observed in favour of students with higher levels of socioeconomic status, home educational resources, and books in the home. Students living in lone-parent households did significantly less well than students living in dual-parent households. Similarly, students who were absent from school on three or more days in the two weeks prior to the PISA assessment performed more poorly than students attending school more regularly. As in mathematics, students in Third year significantly outperformed students in Second year, but did less well than students in Fourth (Transition) and Fifth years. As in mathematics, reading and science, students who studied science for the Junior Certificate Examination outperformed students who did not, providing further evidence that students who do not take science are weaker in a range of areas than students who do take science. Students' reports of both their perceived self-efficacy in mathematics and anxiety about mathematics were associated with performance on problem solving. Students who were most confident in their ability to solve a range of specific mathematics tasks had a higher mean score on problem solving than students with lower levels of confidence. Students with high levels of anxiety about mathematics did less well on problem solving than students who were less confident.

Students attending secondary schools outperformed students in community/ comprehensive schools in problem solving by just over one-fifth of a standard deviation (16.4 points), and students in vocational schools by almost one-half (38.1 points). Students attending schools with the lowest proportions of Junior Certificate Examination fee waiver entitlement achieved a mean score in problem solving that was seven-tenths of a standard deviation (or 55.3 points) higher than that of students attending schools with the highest proportions. As in mathematics, students in schools with high (positive) disciplinary climate levels (as reported by the students themselves) outperformed students in schools with lower levels.

Student-level variables that correlate moderately with problem solving achievement include socioeconomic status (.34) and number of books in the home (.35), while home educational resources (.22) has a somewhat weaker relationship. The correlation between socioeconomic status and problem solving (.34) is about the same as that between socioeconomic status and mathematics (.32). The moderate negative correlation between anxiety towards mathematics and performance in problem solving−(33) indicates that students with higher levels of anxiety tend to perform less well in problem solving (as well as in mathematics) than students with lower levels. The correlation of .31 between fee waiver entitlement at the school level and performance on problem solving provides further support for the use of fee waiver data as a measure of school-level socioeconomic status. The correlation between school-level disciplinary climate in mathematics classes and problem solving is in the weak to moderate range (.19).

In general, the relationships of achievement in problem solving for a variety of background variables are very similar to those for mathematics, reading and science.

8

Conclusions and Implications

CONCLUSIONS

In drawing conclusions about the performance of students in PISA 2003 assessment, it should be borne in mind that PISA is not intended to be a measure of the performance of students on the material they have been taught in school. Rather, the emphasis is on assessing a set of skills in each domain that have been identified by international experts as being important for students' future lives, including their participation in education and in society. It has also been argued that achievement levels in a country are predictive of the country's future economic performance (OECD, 1998), though evidence to support this assertion is outside the scope of PISA.

It should also be noted that the PISA assessments, and their underlying frameworks, represent a particular perspective on what students should be able to do at age 15, and that there are other points of view that should also be considered. Ultimately, policy makers in the areas of curriculum and assessment, as well as schools and teachers, will have to decide on the relevance of the outcomes of PISA for the conduct of learning in schools.

Performance on Combined Mathematics

Ireland's overall performance in PISA 2003 mathematics – 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. Ireland's overall performance in PISA mathematics is consistent with that of Irish post-primary students in earlier international assessments of mathematics. In the 1995 Third International Mathematics and Science Study, for example, Irish students in Second year also achieved a mean score that was not significantly different from the OECD or international country average scores for that study. The average performance of students in Ireland in PISA 2003 mathematics contrasts with their performance in reading and science, where mean scores were above the corresponding OECD country averages.

There are several reasons why performance in mathematics may lag behind that in reading and science. One relates to the relatively poor match between the content/processes of PISA and the Junior Certificate syllabus/examination. Indeed, 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, characterised 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.

A comparison of the performance of Irish students in 2000 and 2003 reveals no significant differences in mean scores in mathematics or in the scores of students at key markers, including the 10th and 90th percentiles. Some might have expected the implementation of the revised Junior Certificate mathematics syllabus (implemented in 2000, for first examination in 2003) to have impacted positively on the performance of Irish students in PISA 2003. However, it was noted that the limited remit of the revision of the syllabus obviated any substantial changes. Hence, it may be that the revised syllabus is not sufficiently different from its predecessor to give rise to a significant change in performance on the mathematics assessed by PISA. Indeed, the ratings of mathematics items common to PISA 2000 and 2003 (which were conducted in 2001 and 2004 respectively) indicated no substantive differences in terms of links between the items and the Junior Certificate syllabus/examination. Second, it may be too soon in the life span of the revised syllabus to expect an immediate impact on student performance. Third, content and teaching methodology may not have changed sufficiently to effect change in the competencies assessed in PISA. Fourth, the Junior Certificate mathematics examination may not have changed in ways that require teachers and students to engage in new approaches to teaching and learning that would impact on performance in assessments such as PISA. Fifth, students taking PISA in 2003 would not have experienced the potential benefits of the new Primary School Mathematics Curriculum (NCCA, 1999), which was not implemented in schools until 2001 at the earliest.

It is noteworthy that the standard deviation associated with Ireland's mean score in mathematics is low relative to other OECD countries (only Finland has a smaller standard deviation). This points to a greater uniformity in achievement in Ireland than in other countries, and perhaps affirms efforts to provide an academic curriculum in mathematics to all students at Junior Cycle level. It is also noteworthy that between-school variance in mathematics achievement in Ireland is low (16.7% compared to an OECD average of 32.7%), indicating that schools are broadly similar to one another in mathematics achievement.

Performance on the Mathematics Subscales

An examination of the performance of Irish students on the four PISA 2003 mathematics subscales points to areas of strength as well as areas of weakness. Strengths lie in the areas of Uncertainty, and, to a lesser extent, Change & Relationships, while the areas of Quantity and Space & Shape are weaker. For Uncertainty and Change & Relationships, mean performance was above the corresponding OECD country average score. Mean performance on the Quantity subscale was not significantly different from the OECD country average, while performance on Space & Shape scale was significantly lower.

The comparatively poor performance of Irish students on Space & Shape was not entirely unexpected. In IAEP 2 in 1991, Irish 13-year olds achieved a mean score in Geometry (a comparable topic area) that was significantly lower than the international average. In TIMSS in 1995, Irish students in Grade 8 (Second year) performed significantly less well on Geometry than they did on the test as a whole. However, the interpretation of the performance of Irish students on Space & Shape in PISA is complicated by the fact that the curriculum-rating project (summarised in Chapter 6) found that none of the PISA Space & Shape items was represented in the Geometry component of the current Junior Certificate syllabus/examination. It was noted that many of the Space & Shape items appear to assess aspects of visual geometry, an area that does not receive much emphasis on the Junior Certificate mathematics syllabus. At Higher and Ordinary levels, concepts underpinning 13 of the 20 PISA Space & Shape items were found to be located in the Junior Certificate syllabus area of Applied Arithmetic & Measure (or 10 of these items, in the case of Foundation level). These findings suggest that any future review of Junior Certificate mathematics syllabus should consider the appropriateness of the current emphasis on teaching and testing abstract and formal knowledge and procedures, compared to a more visually-based approach, such as that embedded in PISA Space & Shape, which includes a greater emphasis on three-dimensional shapes.

Irish students in post-primary schools have done particularly well in some aspects of Quantity in earlier international studies. For example, 13-year olds achieved a mean score above the international average on Number & Operations (a topic area similar to Quantity) in IAEP II, and students in Second year performed better on Fractions & Proportionality (also comparable to aspects of Quantity) than on the test as a whole in TIMSS 1995. In PISA 2003, however, performance on the Quantity subscale was close to the OECD country average. One reason for this may be that, unlike earlier assessments such as TIMSS, items dealing with Quantity in PISA were presented in applied contexts, in line with the philosophy underpinning Realistic Mathematics Education (RME). The assertion that the PISA Quantity items are heavily embedded in applied settings is supported by the findings arising from the curriculum-rating project (Chapter 6) that about one-fifth to one-third of Quantity items were not represented on the Junior Certificate syllabus/examination at all, and that three-tenths to just under one-half were located in the area of Applied Arithmetic & Measure.

The strong performance of Irish students on the Uncertainty scale is interesting, given that one of its components, Probability, does not appear on the Junior Certificate mathematics syllabus. Indeed, the curriculum ratings reported in Chapter 6 show that most Junior Cycle students would be expected to be unfamiliar with 45.0 to 70.0% of PISA Uncertainty items, depending on the level of the syllabus that the students studied. Similarly, in the TIMSS 1995 study, Irish students in Second year achieved a mean score on Data Representation, Analysis & Probability that was significantly higher than their performance on the test as a whole. It may be that Irish students have developed an understanding of aspects of probability and statistics in contexts other than formal mathematics lessons. It may also be the case that, on PISA, where Uncertainty items were presented in concrete and applied settings, Irish students were prepared to take greater risks in attempting to solve problems, than in areas such as Quantity where they would be expected to be familiar with at least some of the underlying concepts.

The above-average performance of Irish students on the PISA Change & Relationships subscale is also difficult to interpret. It arises in part from the fact that the Change and Relationship items in PISA map onto several content areas on the Junior Certificate syllabus, including number systems, applied arithmetic & measure, algebra, statistics, and functions & graphs.

Finally, in considering the performance of students on the PISA subscales, it should be acknowledged that there are aspects of the Junior Certificate mathematics syllabus which PISA did not assess, including sets, trigonometry, and aspects of algebra and geometry.

Performance in Reading

Irish students performed well on reading literacy in PISA 2003, with a ranking of 6th of 29 OECD countries, and 7th of 40 participating countries. This is broadly in line with the strong performance recorded in PISA 2000, when reading literacy was a major domain. Caution should be exercised in interpreting the difference in mean achievement for Irish students between 2000 and 2003 (Irish students had a mean score that was one-tenth of a standard deviation lower in 2003). One hypothesis which can be considered is that the lower performance in 2003 can be attributed to the weaker performance of Irish students at the 75th, 90th and 95th percentiles in particular, where scores were also significantly lower. The uncertainty surrounding the importance of these differences (which also emerged for reading in several other countries, including Canada, Finland and Denmark) suggests that it is necessary to await the outcomes of PISA 2006 before any definite trend in performance can be confirmed. In the meantime, however, some attention should be given to ways in which further improvements in reading can be achieved. The performance of higher-achieving students in particular (considered below) is worthy of further attention.

The relatively poor performance in PISA 2003 of some students taking the Junior Certificate English Examination at Ordinary level is a matter of concern, and merits further analysis. About 27% of Ordinary-level students performed at Level 1 or lower on the reading proficiency scales. Given that performance at or below Level 1 is viewed by the OECD (2004b) as inadequate to meet the future needs of students, and that very few students taking Junior Certificate ordinary level English achieve a grade that is lower than D (just 1.0% in 2003), it would appear that some students who 'pass' Junior Certificate English have weak reading skills (as measured by PISA). This may or may not be a problem depending on one's views of the Junior Certificate Examination and PISA, and whether one takes obvious differences between the two assessments into account (e.g., PISA does not assess essay writing, or the ability to respond to prescribed texts).

Performance in Science

The performance of Irish students in PISA 2003 science was significantly above the OECD country average. There were no changes in mean performance or in the performance of students at key markers since 2000. Given recent concerns over participation in science subjects, particularly at Senior cycle and in third-level institutions (e.g., Task Force on the Physical Sciences, 2002), and performance levels in science in the Certificate examinations, educators hoping for signs of improvement may be disappointed.

It should be noted, however, that science constituted a minor domain in both 2000 and 2003. This means that the assessment tapped into a relatively narrow range of scientific knowledge and concepts, with a particular emphasis on biological and Earth sciences. The new framework for the 2006 assessment, in which students will be assessed on a broader range of scientific concepts and processes, may provide stronger insights into performance of students in Ireland and in other countries.

One development that may be expected to impact on the performance of Irish students in science is the implementation of the new Junior Certificate science syllabus in 2003 (for first examination in 2006). The aims of the new syllabus are more similar to those of PISA, with an emphasis both on knowledge of science and knowledge about science (NCCA, 2002). Moreover, the process of engaging in practical experiments (as part of the Junior Certificate science examination) has the potential to lead to enhanced understanding of scientific concepts. However, the impact of this and of other changes in the syllabus will need to be carefully monitored in both examination and non-examination contexts, to ensure that the intended outcomes are achieved. Such monitoring seems particularly important in the early years of implementation.

A related development, the teaching of the new science syllabus in primary schools since 2003, may also be expected to enhance the performance of students in post-primary schools in the longer term. Again, however, it would be important to monitor the development of scientific knowledge among students in primary schools to ascertain the extent to which intended knowledge and attitudes are acquired.

As in PISA 2000, the current study found that students who do not take science in the Junior Certificate examination (about 10% of 15-year-olds) perform less well in PISA science than students who take the subject. The finding was confirmed in the model of science presented in Chapter 5 in which a negative parameter for the non-study of science was observed, even after adjusting for the effects of a range of school- and student-level variables. The finding that students who do not study science are not significantly different in mean performance on PISA science from students taking the subject at Ordinary level suggests the need for a closer analysis of the performance of students at Ordinary level, particularly lower performers.

Cross-Curricular Problem Solving

Particular care is needed in interpreting the results of the assessment of cross-curricular problem solving, in which students in Ireland achieved a mean score that is not significantly different from the OECD average. On the surface, this assessment appears to test students' ability to solve problems in a range of contexts that are separate from reading, mathematics and science. However, the strong correlations in Ireland between performance on problem solving and performance in the other domains suggests considerable overlap between the assessments, which may be explained by a core set of processes required to solve problems across domains.

Whether or not problem-solving skills can be assessed in large-scale studies such as PISA independently of specific content areas such as mathematics and science is clearly a question that requires further research. Work conducted as part of PISA 2000 in Germany suggests that computer-based assessment of problem solving skills may be one way forward in that it allows students to deal with multiple sources of input, and allows for recursive processing that is not possible in paper-and-pencil assessment contexts (Klieme, 2004). The PISA 2003 assessment of cross-curricular problem solving should be viewed as an initial attempt that requires considerable refinement.

Performance of High and Low Achievers

The scores of lower-achieving students in Ireland – those at the 5th, 10th, and 25th percentiles on combined mathematics, reading, science, and problem solving – were above the corresponding OECD country average scores. This may be a function of the relatively small number of non-national students, and/or students who speak a language other than the language of instruction in Irish schools. This is not to say that such students are by default low achievers; rather, their lack of familiarity with the language(s) of instruction may act as a barrier to maximising their potential.

Given that all three models reported in Chapter 5 showed negative effects on achievement for pupils in schools with the highest levels of deprivation (based on the percentages of students entitled to a Junior Certificate Examination fee waiver), efforts to support and improve the achievements of students in disadvantaged circumstances should continue.

The low performance of higher-achieving students in Ireland must be regarded as a matter of concern. In PISA 2003 combined mathematics, for example, even though the average performance of Irish students does not differ significantly from the OECD country average, Irish scores at the 75th, 90th and 95th percentile ranks are significantly lower than the corresponding OECD country averages. This, coupled with the relatively small percentage of students achieving Levels 5 and 6 on the combined mathematics proficiency scales (11.3%, compared to an OECD average of 14.6%, and 23.4% in Finland, the highest performing country), points to underperformance among higher achievers in Ireland. A similar pattern is evident in science and problem solving, while the performance in reading of students scoring at these percentile ranks was significantly lower in 2003 than in 2000.

Among the issues that might be considered in examining the relatively low performance of high achievers are the effects of syllabi and examinations on performance. In mathematics, for example, it may be the case that the syllabi and indeed examinations provide insufficient scope for students (including higher achievers) to establish relational understanding of content, as teachers and students strive to cover as much content as possible in preparation for examinations. Using the terminology of the PISA framework, it may be that students do not get sufficient opportunities to engage in horizontal (rather than vertical) mathematisation, where they have opportunities to use their mathematical knowledge in a range of applied settings, and in so doing, consolidate and extend that knowledge. There is support for this view when one considers that the Junior Certificate mathematics syllabus includes complex algebraic techniques and geometric proofs that are not tested by PISA. Again, these may well be challenging to students, but they reflect a focus on vertical mathematisation, rather than the horizontal mathematisation emphasised by PISA.

Alternatively, teachers may attend to the needs of weaker students in their classes to such an extent that there is insufficient time available to extend the highest achievers. One possible way forward would be to consider whether the upper range of some higher-level examinations (e.g., English) could be extended to accommodate (and motivate) the highest-achieving students. Another would be to offer a series of additional modules to higher-achieving students so that they have opportunities to reach their potential, perhaps outside the constraints of examination content.

Contributions of Student-Level Variables to Achievement

The multilevel models of achievement in combined mathematics, reading literacy and science revealed significant gender differences in the presence of other school- and student-level variables, although the strength and direction of these differences varied according to the domain. Differences are of a similar magnitude for combined mathematics (three-tenths of a standard deviation in favour of males) and reading (close to one-quarter of a standard deviation in favour of females), while the adjusted value for science is smaller (just under one-tenth of a standard deviation in favour of males), having changed from non-significant when originally fitted separately, to statistically significant.

Although not investigated in the context of a multilevel model of achievement, the particularly large gender difference in favour of male students (close to one-third of a standard deviation) on the Space & Shape subscale is worthy of further examination in the context of examining overall gender differences in PISA mathematics. It might be argued, for example, that although gender differences in favour of male students were observed on all four PISA mathematics subdomains, the gender difference in Space & Shape contributes disproportionately to the size of the overall difference. The gender differences in favour of females in Junior Certificate mathematics, reported in Chapter 6, contrast with those observed in PISA mathematics, suggesting that male students are at an advantage on assessments such as PISA, which focus less on content covered in mathematics classes, and more on application of mathematics knowledge and problem solving in real-life contexts. However, it is also noteworthy that, despite strong correlations between students' scores on the PISA combined mathematics and cross-curricular problem solving, the unadjusted mean score difference between males and females on the latter was negligible and not statistically significant.

The gender difference in reading literacy is consistent with that found in PISA 2000 and with the gender differences in Junior Certificate English. Moreover, the finding in both PISA 2000 and PISA 2003 that significantly more males are at Levels 1 and 2, and significantly more females at Levels 4 and 5 on the combined reading proficiency scale, suggests that ongoing, early targeted intervention for boys with weak reading skills is needed if greater equity in performance in reading is to be achieved.

The adjusted gender difference for science, which emerged in the multilevel model of performance reported in Chapter 5, indicates that a difference in achievement in favour of male students emerges as one controls for the effects of a range of relevant variables, including individual socioeconomic status and whether or not a student studied science for the Junior Certificate. It would be worthwhile, in future research, to explore which particular set of variables is associated with the gender difference, especially in the context of the slightly lower take-up of science at Junior Certificate by female students compared to males, and the superior performance of females on the Junior Certificate Examination in science.

The finding that student SES is related to achievement is not new. Related factors, such as books in the home and home educational resources, retained significant associations with performance when included in models of achievement. These would appear to measure aspects of home climate that are supportive of students' school experiences.

It was noted that Irish students from lone parent families were particularly at risk of low achievement, relative to similar students in other participating countries, and an achievement deficit (relative to students in dual parent families), in the order of one-sixth of a standard deviation, remains in the presence of the other school- and student-level variables.

Contributions of School-Level Variables to Achievement

When student-level variables were fitted alone (without the school-level variables) they explain a substantial percentage of the variance in achievement between schools. This indicates that considerable variation in school performance due to differences in student composition. However, school-level variables (SES and disciplinary climate) explained significant amounts of between-school variation, over and above those at the student level in the three final models presented in this report. In particular, this illustrates the contextual effect of the school-level SES-related variable over and above the individual student measures used in the model.

The impact of SES at the school level is confirmed, with the addition (in 2003) of a useful school-level deprivation context measure that is easily collected and is policy relevant: percent of students entitled to Junior Certificate Examination fee waiver. The variable captures the continuous nature of SES density in the school context. In contrast to PISA 2000, school sector does not appear in any of the final models for 2003, indicating that the combination of other variables in the models explain the achievement differences across secondary, vocational, and community/comprehensive schools.

Students in schools with a high positive average disciplinary climate in their mathematics class (as perceived by the students themselves) had higher expected mean scores in all assessment domains than students in schools with medium and low levels. Further, school-level socioeconomic status and disciplinary climate are only weakly related to one another. Further development of the disciplinary climate measure could yield promising results.

IMPLICATIONS

In this section, implications arising from the outcomes of PISA 2003 in Ireland are presented. Almost half relate specifically to mathematics. The others relate to reading, science, cross-curricular problem solving, as well as school- and student-level variables associated with achievement.

Mathematics

1. *Overall performance in mathematics.* The finding that the overall mean score of students in Ireland, although not significantly different from the OECD country average, is lower than the mean score of students in a number of countries, prompts the question: Should we be satisfied with current standards and do they meet current and future needs? In considering this, it should be recognised that, although the stated aims of PISA

mathematics and of the Junior Certificate mathematics syllabus are broadly similar, there are substantial differences in their underlying philosophies[59], in mathematical content, in assessment, and in implications for teaching and learning.

2. *Performance on the mathematics subdomains (subscales).* In all four mathematics subdomains, mean scores for Ireland were well below those of the highest-scoring countries. The weakest subdomain was Space & Shape, where performance was significantly below the OECD country average. The performance of students in Space & Shape may reflect differences between the focus on spatial reasoning and visual geometry in PISA, and the emphasis on deductive, logical geometry in the Junior Certificate syllabus/examination. Indeed, none of the Space & Shape items in PISA was identified by individuals experienced in the mathematics curriculum in Ireland as an element of the syllabus in geometry (see 8 and 9, below).

3. *Distribution of achievement in mathematics.* The standard deviation around the mean score of students in Ireland on combined mathematics (85.3) was among the lowest in the OECD – well below the country average standard deviation (100.0). A low standard deviation indicates a narrow dispersion of achievement scores, and can be interpreted as evidence of an equitable distribution of achievement outcomes in a country. Countries with narrow dispersions may of course have high or low overall achievement.

4. *Performance of high achievers in mathematics.* The relatively low performance of higher-achieving students in mathematics in Ireland is noteworthy, as students scoring at the 75th, 90th and 95th percentiles achieved scores that were significantly below the corresponding OECD country average scores. This suggests that any forthcoming review of mathematics at post-primary level should include a consideration of these outcomes, with a view to identifying ways in which performance of high achievers might be enhanced. It is likely that enhancement could be enabled within the structure of existing syllabi/examinations, by providing opportunities for higher-achieving students to demonstrate their abilities on problems that are more strongly embedded in real-life contexts, and recognising their achievements on such work. In addition, wider access to opportunities at Senior Cycle level that would engage higher-achieving students in more in-depth application and problem solving around mathematical concepts they have acquired, while rewarding them academically for doing so, might merit consideration.

5. *Performance of low achievers in mathematics.* The observations that 16.8% of all students in Ireland, and 21.9% of students taking the Junior Certificate mathematics examination at Ordinary level, score at or below Level 1 on the PISA mathematics proficiency scale, are obvious matters of concern. It would be worth investigating if the performance of lower-achieving students could be enhanced by providing them with learning experiences in which conceptual understanding is more strongly emphasised, and relationships between mathematics knowledge and real-life applications are more obvious.

6. *Gender differences in mathematics.* Male students in Ireland achieved a mean score on combined mathematics that was significantly higher (by 14.8 points, or one-sixth of a standard deviation) than that of females. Gender differences favouring males were also observed on the four mathematics subdomains, with the largest occurring for Space & Shape scale (just over a quarter of a standard deviation). These differences contrast with those found for all students taking the Junior Certificate mathematics examination in 2003, where females outperform males by about one-half of a grade (one-sixth of a standard deviation). Further investigation of the differences might consider associations between gender and performance on different problem types (including those that call for

[59] The framework for PISA mathematics is grounded in the Realistic Mathematics Education movement.

spatial reasoning), as well as the differing risk-levels that male and female students may adopt in responding to PISA-type problems in a low-stakes assessment context.

7. *Self-efficacy in mathematics and anxiety about mathematics.* Similar to other participating countries, male students in Ireland reported higher self-efficacy in mathematics (confidence in their ability to solve mathematics problems) and lower anxiety about mathematics than females. However, interpretation of the associations of these variables with achievement in mathematics warrants care, since self-efficacy and anxiety may affect and be affected by students' current mathematics performance.

8. *Concepts underlying PISA and Junior Certificate Mathematics*. The test curriculum rating project conducted in conjunction with the Irish analysis of PISA 2003 showed that students in Ireland were expected to be very familiar or familiar with the mathematical concepts underlying between 48.3% and 69.4% of PISA items (depending on the syllabus level taken). This suggests that any future review of mathematics at post-primary level should consider if important mathematical content is absent. Again, any debate around the differences between PISA and Junior Certificate mathematics would need to recognise that important elements of the Junior Certificate mathematics syllabus are not assessed by PISA (e.g., sets, geometry and trigonometry), that some PISA concepts (e.g., probability) only appear on mathematics syllabi at Senior Cycle level and that students may acquire mathematical concepts outside mathematics classes.

9. *Context in PISA and Junior Certificate mathematics items.* The test curriculum rating project showed that students in Ireland were expected to be very familiar or familiar with just 20.0% to 34.1% of the contexts in which PISA mathematics items were embedded. Any forthcoming review of mathematics at post-primary level might consider whether greater attention should be paid to interpreting and solving mathematics problems embedded in realistic contexts across different mathematics topics (described as 'horizontal mathematisation' in Realistic Mathematics Education), and how the teaching and assessment of mathematics might be affected by such a change.

10. *Between-school variation in achievement in mathematics.* In Ireland, just 16.7% of the variation in mathematics achievement is attributable to differences between schools. This indicates that, relative to many other countries, schools tend to be more alike in mathematics achievement. Most of the variation in achievement in Ireland is within schools – between classes and students.

11. *Implementation of the revised Junior Certificate mathematics syllabus.* The finding that the mean mathematics performance of students in Ireland in 2000 and 2003 did not differ significantly may suggest that the revised Junior Certificate mathematics syllabus (introduced in 2000, for first examination in 2003) has not yet had the expected impact. A comparison of the PISA mathematics items and items presented on the revised Junior Certificate mathematics examination papers indicate large differences in the manner in which mathematics is presented; perhaps the intended increase in emphasis on relational understanding and problem solving is not yet sufficiently evident in what is actually assessed in Junior Certificate mathematics. However, actual implementation may also take longer to have a significant impact on achievement.

Reading

12. *Overall performance in reading*. Just three countries in PISA 2003 had significantly higher mean scores in reading than Ireland, while Ireland's mean score was well above the OECD country average. This indicates that, despite having a slightly lower mean score than in 2000, Ireland has maintained its status as one of the highest-performing countries in reading literacy.

13. *Performance in reading in 2000 and 2003.* Reading is the only PISA domain in which a difference in the performance of students in Ireland was observed between 2000 and 2003. The mean score was significantly lower in 2003, while the scores at the 75th, 90th and 95th percentile points in Ireland were also significantly lower. No differences were observed in performance at the 5th, 10th, or 25th percentiles.

14. *Gender differences in reading.* Female students in Ireland achieved a mean score in PISA reading that was 29.0 points (one-third of a standard deviation) higher than that of males. The difference in favour of female students in PISA on the Junior Certificate Examination was almost one grade point (two-fifths of a standard deviation). There is a need to further examine the combination of variables corresponding to gender differences in English/reading on both PISA and the Junior Certificate, including socioeconomic status, motivation to read, and reading habits, before interventions can be designed to align the performance of males and females (a challenge faced by most countries participating in PISA).

15. *Low achievers in reading.* In both 2000 and 2003, approximately 11% of students in Ireland achieved at or below Level 1 in reading. Although lower than the OECD country average percentage in both years, the findings point to a lack of progress in reducing the proportion of students in schools who have reading difficulties.

Science

16. *Overall performance on science.* The mean score of Irish students in PISA 2003 science was significantly higher than the OECD country average, though the gap between Ireland's mean score and those of the highest-scoring countries was considerable, indicating room for improvement. More detailed information on performance in science will become available in 2006, when science becomes a major assessment domain in PISA, and more definite strategies for improvement can be formulated, if required.

17. *Study of science.* As in PISA 2000, students in PISA 2003 who reported that they did not study science as a subject for the Junior Certificate Examination (5.2% of males; 14.6% of females; 9.9% of all students) achieved a mean score that was lower by 69.7 points (four-fifths of a standard deviation) than that of students who took science as a subject. Although students who did not take science also performed less well in mathematics and reading than those who did take the subject, it would seem important that all First-year students are well informed about the benefits of choosing science as a subject (for example, by providing short 'taster' courses in science). The range of reasons why students opt not to study science, as well as science topics that might be most relevant to these students' needs, merit further consideration. In any event, it would seem important to make efforts to develop the scientific knowledge of all Junior Cycle students, although it may not be necessary for all of them to take the science in the Junior Certificate Examination to accomplish this.

18. *Gender differences in science.* Whereas the mean scores of male and female students in PISA science are not significantly different when looked at in isolation, a small but significant difference of 7.6 points (just under one-tenth of a standard deviation) in favour of male students was observed when the contributions of a range of factors were considered simultaneously. The precise combination of variables that contribute to this significant gender difference is not clear. This finding differs from the finding relating to science in the Junior Certificate Examination. Females taking the examination in 2003 scored about half a grade point (a quarter of a standard deviation) higher. This difference should be examined further to ascertain if it arises for methodological (e.g., the use of simulated scores for students who did not take the PISA science assessment) or other reasons.

19. *Performance of high achievers in science.* Although overall performance on PISA science is significantly above the OECD country average, Irish students scoring at the 90th and 95th percentiles achieved scores that are significantly below the corresponding OECD country averages. This is consistent with the comparatively low performance of higher achievers in mathematics, and with the decline in performance among higher achievers in reading between 2000 and 2003. It further reinforces the need to investigate ways to extend the knowledge and skills of the highest performers in schools in Ireland.

Cross-Curricular Problem Solving

20. *Performance on cross-curricular problem solving*. The mean score on cross-curricular problem solving was not significantly different from the OECD country average. However, as in mathematics, the scores of Irish students at the 75th, 90th, and 95th percentiles were below the corresponding OECD country average scores, while those at the 5th, 10th and 25th percentiles were above the corresponding OECD average scores.

21. *Problem solving and combined mathematics.* There are quite strong correlations among the four domains assessed in PISA 2003. The correlation between problem solving and combined mathematics for students in Ireland is .90, indicating that a large amount of variation in each assessment is shared. The strong correlation between mathematics and problem solving suggests that, despite differences in content, they may share some of the same underlying cognitive processes. There may be some value in investigating which specific processes are common to both assessments, and what factors might contribute to the development of these.

School-Level Variables

22. *School socioeconomic status and performance*. Multilevel models of achievement in mathematics, reading and science highlight the contributions of individual- and school-level socioeconomic status to achievement. The level of disadvantage associated with the school that a student attends (based on the percentage of students who are entitled to a Junior Certificate Examination fee waiver) had a significant association with achievement in combined mathematics, reading and science, even after adjustments had been made for other school and student variables, including student socioeconomic status. This finding justifies current efforts to target resources on schools with large numbers of disadvantaged students. It also supports the use of fee-waiver data as a measure of school-level deprivation as an alternative to the dichotomous designated disadvantaged/not designated variable used in other analyses. Moreover, the correlation between the fee-waiver data variable and an international combined measure of economic, social and cultural status (aggregated to the school level) is strong (–.81), further supporting the validity of the use of the fee-waiver measure as an indicator of school-level disadvantage.

23. *School disciplinary climate and performance*. After adjusting for the effects of other school- and student-level variables, students in schools with a high positive average disciplinary climate in their mathematics classes (as rated by the students themselves) had higher expected mean scores in all assessment domains than students in schools with medium and low levels. Further, the correlation between school-level socioeconomic status and disciplinary climate is weak, indicating that climate is independent of school socioeconomic status. These findings suggest that the nature of disciplinary climate might be examined further, to ascertain if there are ways in which learning environments could be better structured to support student learning. It would also be useful to ascertain why school-level disciplinary climate in mathematics also explains achievement in reading and science (perhaps by being a good proxy for school disciplinary climate in general).

Student-Level Variables

24. *Lone-parent status and performance*. Even after adjusting for other relevant variables, lone-parent status made a negative contribution to students' achievement in combined mathematics, reading and science (e.g., 15.7 points or one-sixth of a standard deviation in mathematics). This indicates that students in lone-parent families are particularly at risk of low achievement, and might be the focus of interventions that take their needs and the needs of their parents into account. The finding that the achievement gap in mathematics between students in lone-parent and dual-parent families is larger than that of Ireland in only two countries in PISA 2003 suggests that students in lone-parent families in Ireland are particularly at risk in comparative terms.

25. *Absenteeism and performance.* Attendance at school was significantly associated with achievement in reading and science (although the measure was limited to the two weeks prior to the PISA assessment, and reasons for the absences are unknown). In mathematics, there was an interaction between attendance and the index of books in the home. Students who were absent for three or more days and who had few books in the home had the lowest fitted scores. These findings underline the value of supporting regular attendance of students in school, particularly those from backgrounds where literacy activities may not be emphasised.

26. *Home educational processes and performance.* In addition to confirming the effects of student socioeconomic status on achievement, the multilevel models of achievement confirm the contribution of the number of books in a student's home and the availability of home educational resources to achievement in reading, mathematics and science.

LOOKING AHEAD TO PISA 2006

The next PISA assessment is scheduled for 2006. It will offer an opportunity to extend our knowledge about the performance of 15-year-olds in Ireland and in other countries. First, it will offer a second opportunity to examine changes in achievement, allowing more confident interpretations to be made, as methodologies become more refined. Second, science will become the major focus of the assessment, allowing an in-depth examination of students' achievements across a wider range of content areas than has been possible in either 2000 or 2003, together with the opportunity to describe student achievements along proficiency levels similar to those which already exist for mathematics and reading. Third, PISA 2006 will coincide with the first examination of the revised Junior Certificate science syllabus and will be particularly useful in aiding the exploration of issues surrounding its implementation. To this end, a detailed comparison of the revised curriculum and PISA, similar to the one that was carried out for mathematics, is planned, while it is also planned to ask teachers about syllabus implementation. Fourth, an international optional computer-based assessment of science may be implemented in some countries, including Ireland. The field trial of the computer-based assessment takes place alongside the regular PISA activities in 13 countries in the spring of 2005. It is envisaged that the computer-based assessment will add value to the paper-and-pencil assessment by allowing for the assessment of skills in a more dynamic and interactive environment. However, the success of the computer-based assessment and the form it will take for 2006 will only be decided following the analysis of the pilot data collected in the field trial. Finally, the inclusion of attitudinal items, which assess student interest in science, their support for scientific research, and their sense of responsibility towards environmental and other concerns in the test booklets alongside the test items, will represent a novel way of assessing attitudes to science. However, as with the computer-based assessment, the final design of these items has yet to be decided.

References

Adams, R.J., & Wu, M.L. (2002). *PISA 2000 technical report.* Paris: Organisation for Economic Co-operation and Development.

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

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

Beaton, A.E., Mullis, V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.L., & Smith, T.A. (1996a). *Mathematics achievement in the middle-school years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Beaton, A.E., Martin, M.O Mullis, I.V., Gonzalez, E.J., Smith, T.A., & Kelly, D.L. (1996b). *Science achievement in the middle-school years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18 (1),* 32-42.

Carey, M. (1990). Gender differences in attitudes and achievement in mathematics: A study of first year students in Irish post-primary schools. *Unpublished M.Ed. Thesis.* Ireland: University of Dublin (Trinity College).

Cook, R.D., & Weisberg, S. (1999). *Applied regression including computing and graphics.* New York: Wiley.

Cosgrove, J., Sofroniou, N., Kelly, A., & Shiel, G. (2003). *A teacher's guide to the reading literacy achievements of Irish 15-year olds.* Dublin: Educational Research Centre.

Cosgrove, J., Shiel, G., & Kennedy, D. (2002). The performance of Irish students in scientific literacy in the Programme for International Student Assessment. *Irish Journal of Education, 33,* 53-70.

Csapó, B. (1997). The development of inductive reasoning: Cross-sectional assessments in an educational context. *International Journal of Behavioral Development, 20,* 609-626.

Department of Education and Science (2001). *Annual statistical report 2000/2001.* Dublin: Author.

Department of Education and Science/National Council for Curriculum and Assessment. (1999). *Primary school curriculum: Mathematics.* Dublin: Stationery Office.

Department of Education and Science/National Council for Curriculum and Assessment. (2000). *Mathematics syllabus: Higher, Ordinary, and Foundation Level.* Dublin: Stationery Office.

Dunn, O.J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, *56*, 52-64.

Eivers, E., Shiel, G., & Shortt, F. (2004). *Reading literacy in disadvantaged primary schools.* Dublin: Educational Research Centre.

Elley, W.B. (1994). *The IEA Study of Reading Literacy: Achievement and instruction in thirty-two school systems.* Oxford: Pergamon.

Eurydice (2004). *Integrating immigrant children into schools in Europe: Ireland.* Brussels: Eurydice European Unit. Retrieved October 10 2004 from http://www.eurydice.org/Documents/Mig/en/frameset_immigrant.html

Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht: Kluwer Academic.

Freudenthal, H. (1981). Major problems in mathematics education. *Educational Studies in Mathematics, 12,* 133-150.

Ganzeboom, H.B., De Graaf, P., & Treiman, D.J. (with De Leeuw, J.) (1992). A standard international socioeconomic index of occupational status. *Social Science Research*, *21,* 1-56.

Ganzeboom, H.B., & Treiman, D.J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research*, *25,* 201-239.

Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Griffin.

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

Hox, J. (2002). *Multilevel analysis: Techniques and applications.* Mahwah, NJ: Lawrence Earlbaum Associates.

Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. London: Sage.

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

Kirsch, I., de Jong, J., Lafontaine, D., McQueen, J., Mendelovits, J., & Monseur, C. (2002). *Reading for change: Performance and engagement across countries. Results from PISA 2000.* Paris: OECD.

Klare, G.R. (1984). Readability. In P.D. Pearson (Ed.), *Handbook of reading research* (pp. 681-744). London: Longman.

Klieme, E. (2004). Assessment of cross-curricular problem-solving competencies. In J.H. Moskowitz & M. Stephens (Eds.), *Comparing learning outcomes: International assessment and educational policy.* London: Routledge-Falmer.

Lapointe, A.E., Mead, N.A., & Phillips, G.W. (1989). *A world of difference: An international assessment of mathematics and science.* Princeton, NJ: Educational Testing Service.

Lapointe, A.E., Askew, J.M., & Mead, N.A. (1992). *Learning science.* Princeton, NJ: Educational Testing Service.

Lapointe, A.E., Mead, N.A., & Askew, J.M. (1992). *Learning mathematics: The International Assessment of Educational Progress.* Princeton, NJ: Educational Testing Service.

Lindsey, J.K., & Lindsey, P.J. (2001). Detecting covariates with non-random missing values in a survey of primary education in Madagascar. *Journal of the Royal Statistical Society, 164 (Part 2),* 327-338.

Little, R.J., and Schenker, N. (1994). Missing data. In G. Arminger, C.C. Clogg & M.E. Sobel (Eds.), *Handbook for statistical modeling in the social and behavioral sciences* (pp. 39-75). New York: Plenum.

Longford, N. (1993). *Random coefficient models.* Oxford: Clarendon 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.

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., Hickey, B.L., & Murchan, D.P. (1992). The Second International Assessment of Educational Progress: Mathematics and science findings in Ireland. *Irish Journal of Education, 26,* 3-146.

Martin, M.O., & Morgan, M. (1994). Reading literacy in Irish schools: A comparative analysis. *Irish Journal of Education, 28,* 3-101.

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.

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

Mislevy, R.J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika, 56,* 177-196.

Morgan, M., Hickey, B., & Kellaghan, T. (1997). *International Adult Literacy Survey: Results for Ireland. Report to the Minister for Education.* Dublin: Stationery Office.

NCCA (National Council for Curriculum and Assessment). (2002). *Response to PISA 2000.* Retrieved March 14, 2005, from http://www.ncca.ie/j/Pisa2000.pdf

Oldham, E. (1989). Is there an international mathematics curriculum? In B. Greer & G. Mulhern (Eds.), *New directions in mathematics education* (pp. 185-224). London: Routledge.

Oldham, E. (2002). The performance of Irish students in mathematical literacy in he Programme for International Student Assessment (PISA). *Irish Journal of Education, 33,* 31-52.

O'Leary, M., Kellaghan, T., Madaus, G.F., & Beaton, A.E. (2000). Consistency of findings across international surveys of mathematics and science achievement: A comparison of IAEP2 and TIMSS. *Education Policy Analysis Archives, 8,* 43. Retrieved September 1, 2004, from http://epaa.asu.edu/epaa/v8n43/html

OECD (Organisation for Economic Co-operation and Development). (1993). *Education at a glance: OECD indicators*. Paris: Author.

OECD (Organisation for Economic Co-operation and Development). (1997). *Education at a glance: OECD indicators*. Paris: Author.

OECD (Organisation for Economic Co-operation and Development). (1998). *Human capital investment: An international comparison*. Paris: Author.

OECD (Organisation for Economic Co-operation and Development). (1999). *Classifying educational programmes: Manual for ISCED-97 implementation in OECD countries.* Paris: Author.

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

OECD (Organisation for Economic Co-operation and Development). (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills*. Paris: Author.

OECD (Organisation for Economic Co-operation and Development). (2004a). *Education at a glance: OECD indicators*. Paris: Author.

OECD (Organisation for Economic Co-operation and Development). (2004b). *Learning for tomorrow's world - First results from PISA 2003.* Paris: Author.

OECD (Organisation for Economic Co-operation and Development). (2004c). *Problem solving for tomorrow's world. First measures of cross-curricular skills from PISA 2003. Paris: Author.*

OECD (Organisation for Economic Co-operation and Development)/HDRC (Human Resources Development, Canada) (1997). *Literacy skills for the knowledge society: Results from the First International Adult Literacy Survey*. Paris: OECD.

OECD (Organisation for Economic Co-operation and Development)/Statistics Canada (2000). *Literacy in the information age: Final report of the International Adult Literacy Survey.* Paris: OECD.

OECD (Organisation for Economic Co-operation and Development)/UNESCO-UIS (United Nations Educational, Scientific and Cultural Organisation Institute for Statistics). (2003). *Literacy skills for the world of tomorrow. Further results from PISA 2000.* Paris: Authors.

Pinheiro, J.C., & Bates, D.M. (1999). *Mixed-effects models in S and S-Plus.* New York: Springer.

PISA Consortium. (2001). *Standards for PISA 2003.* Melbourne: Australian Council for Educational Research.

PISA Consortium. (2003). *Data adjudication in PISA 2003*. Melbourne: Australian Council for Educational Research.

Polya, G. (1945). *How to solve it.* Princeton, NJ: Princeton University Press.

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., Bryk, A.S., Cheong, Y.F., & Congdon, R.T. (2000). *Hierarchical linear and nonlinear modeling*. IL: Scientific Software International.

Schafer, J.L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Schmidt, W.H., McKnight, C.C., Valverde, G.A., Houang, R.T., & Wiley, D.E. (1997). *Many visions, many aims. Volume I: A cross-national investigation of curricular intentions in school mathematics.* Dordrecht, NL: Kluwer Academic.

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

Shiel, G., & Kelly, D. (2001). *The 1999 national assessment of mathematics achievement.* 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.

Sofroniou, N. (2004, May). *The reading skills of Irish 15-year olds in the OECD Programme for International Assessment (PISA 2000).* Paper presented at the 24th Conference on Applied Statistics in Ireland, Galway.

Sofroniou, N., Cosgrove, J., & Shiel, G. (2002). Using PISA variables to explain performance on Junior Certificate examinations in mathematics and science. *Irish Journal of Education, 33,* 99-124.

Sofroniou, N., Shiel, G., & Cosgrove, J. (2000). Explaining performance on the Junior Certificate examination: A multi-level approach. *Irish Journal of Education, 31,* 25-49.

Sofroniou, N., Shiel, G., & Cosgrove, J. (2002). PISA reading literacy in Ireland: An expanded model exploring attributes of self-regulated learning. *Irish Journal of Education, 33,* 71-98.

State Examinations Commission. (2003). *Junior Certificate Examination 2003: Mathematics: Chief Examiners' reports*. Athlone: Author. Retrieved September from http://www.examinations.ie/main.php?l=en&mc=co&sc=ct

Task Force on the Physical Sciences. (2002). Report and recommendations. Dublin: Department of Education and Science. Retrieved September 2004 from http://www.sciencetaskforce.ie/report/report.pdf

Treffers, A. (1987). *Three dimensions: A model of goal and theory description in mathematics instruction – the Wiskobas project.* Dordrecht: Kluwer Academic Publishers.

Vosniadou, S., & Orthony, A. (1989). *Similarity and analogical reasoning.* New York: Cambridge University Press.

Westat. (2000). *WesVar complex samples 4.2*. Rockville, MD: Author.

Appendix A

PISA Test Design, and Sample Tasks for Mathematics and Problem Solving

A.1 PISA 2003 Rotated Test Design

Table A.1 shows the allocation of clusters of test items to the 13 test booklets used in PISA 2003. Test items were allocated to 13 30-minute blocks or clusters. There are seven mathematics clusters (M_1 – M_7), two science clusters (S_1 – S_2), two problem solving clusters (P_1 - P_2) and two reading clusters (R_1 – R_2). The test design has the following characteristics: each cluster appears four times in the design, once in each position in a booklet. Mathematics items appear in all 13 booklets; some booklets have just one mathematics cluster, others have up to three. All booklets have clusters from at least two assessment domains and one booklet (Booklet 9) has a cluster from each of the four domains.

Table A.1. Test Booklet Design for PISA 2003

Booklet	*Cluster 1 30 mins*	*Cluster 2 30 mins*	*Cluster 3 30 mins*	*Cluster 4 30 mins*
1	M1	M2	M4	R1
2	M2	M3	M5	R2
3	M3	M4	M6	PS1
4	M4	M5	M7	PS2
5	M5	M6	S1	M1
6	M6	M7	S2	M2
7	M7	S1	R1	M3
8	S1	S2	R2	M4
9	S2	R1	PS1	M5
10	R1	R2	PS2	M6
11	R2	PS1	M1	M7
12	PS1	PS2	M2	S1
13	PS2	M1	M3	S2

A.2 Interpreting the Tables in Appendix A

Sample tasks and items for mathematics and problem solving have been adapted from the *PISA 2003 International Report* (OECD, 2004b). The layout has been compacted somewhat and is not identical to that which the students are presented in the test booklets. For each item the information under 'PISA Item Difficulty' gives the IRT scale score that represents the location of the item on the relevant achievement scale (in terms of average performance of a fixed number of pupils drawn from each OECD countries). Each scale has a mean of 500.0 and a standard deviation of 100.0. Item difficulties are also reported in terms of the proficiency levels at which they are located. In cases where an item offers both partial credit (PC) and full credit (FC), item difficulties and proficiency levels for both levels of credit are given.

Additional item statistics are provided in this appendix for the sample items. These include the OECD percentage correct score and the weighted percentage of Irish students providing a correct response, the percentages giving an incorrect response, and the percentages not responding. Again, where appropriate, percentages of students receiving partial and full credit on an item are given.

Example Interpretations

In the sample mathematics item Exchange Rate Question 1 (p. 233), the item difficulty across OECD countries is 406.1 points. This means that the item is located almost a full standard deviation below the OECD mean of 500.0. Further, since the item is between 358.3 and 420.4 score points on the mathematics scale, the item is at proficiency Level 1. The OECD average percentage correct for the item is 79.7%, while, in Ireland, it is slightly higher at 83.2%. The OECD average 'missingness' for the item is 6.6%, while in Ireland, it is 3.5%.

The sample mathematics item Growing Up Question 2 (p. 239) is an item on which both partial and full credit are available, depending on the completeness of a student's answer. The fully-correct version of the item has a difficulty estimate of 525.3, while the partially-correct version has an estimate of 419.3. The full-credit estimate is at proficiency Level 3, while the partial credit version is on the border between Levels 1 and 2. On average across OECD countries, 54.7% of students received full credit for their responses, while in Ireland, 51.5% did so. The corresponding percentages for partial credit were 28.1% and 35.8% respectively. Again, missingness in Ireland was lower than the OECD average level.

Further Sample Tasks

As no new items for reading literacy or science were released following the PISA 2003 assessment, sample items for these domains were drawn from a pool of items released following the 2000 assessment (for sample tasks for reading and science, see OECD, 2001; Shiel et al., 2001; Cosgrove et al., 2003). A more detailed and comprehensive appendix, which includes a larger number of sample items from all four assessment domains, is available at http://www.erc.ie/pisa.

A.3 Sample Tasks

MATHEMATICS

UNIT: EXCHANGE RATE (2003)

Context: *Public.*

Mei-Ling from Singapore was preparing to go to South Africa for 3 months as an exchange student. She needed to change some Singapore dollars (SGD) into South African rand (ZAR).

EXCHANGE RATE QUESTION 1 (Item code: M413Q01)

Domain: *Quantity.* **Item type**: *Short constructed response.*

Mei-Ling found out that the exchange rate between Singapore dollars and South African rand was:
1 SGD = 4.2 ZAR.
Mei-Ling changed 3000 Singapore dollars into South African rand at this exchange rate.
How much money in South African rand did Mei-Ling get?

Key: *Full credit:* 12 600 ZAR (unit not required); *no credit:* Other responses, missing.
Process: *Reproduction.* Understand a simple problem, and link the given information to the required calculation.

PISA Item Difficulty
Scale score
406.1
Level
1

Item statistics	% OECD	% Ireland
Correct	79.7	83.2
Incorrect	13.8	13.4
Missing	6.6	3.5
Total	100	100

EXCHANGE RATE QUESTION 2 (Item code: M413Q02)

Domain: *Quantity.* **Item type**: *Short constructed response.*

On returning to Singapore after 3 months, Mei-Ling had 3 900 ZAR left. She changed this back to Singapore dollars, noting that the exchange rate had changed to: 1 SGD = 4.0 ZAR.
How much money in Singapore dollars did Mei-Ling get?

Key: *Full credit*: 975 SGD (unit not required); *no credit*: Other responses, missing.
Process: *Reproduction.* Understand a simple problem, and decide that division is the right procedure to go with.

PISA Item Difficulty
Scale score
438.8
Level
2

Item statistics	% OECD	% Ireland
Correct	73.9	76.3
Incorrect	17.3	18.2
Missing	8.8	5.5
Total	100	100

EXCHANGE RATE QUESTION 3 (Item code: M413Q03)

Domain: *Quantity.* **Item type**: *Open constructed response.*

During these 3 months the exchange rate had changed from 4.2 to 4.0 ZAR per SGD.
Was it in Mei-Ling's favour that the exchange rate now was 4.0 ZAR instead of 4.2 ZAR, when she changed her South African rand back to Singapore dollars? Give an explanation to support your answer.

Key: *Full credit*: 'Yes', with adequate explanation; *no credit*: 'Yes', with no explanation or with inadequate explanation, other responses, missing.
Process: *Reflection.* Identify the relevant mathematics, reduce the task to a problem within the mathematical world, and construct an explanation of the conclusion.

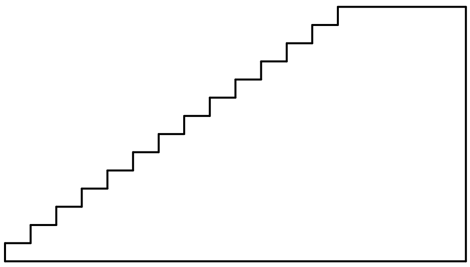
PISA Item Difficulty
Scale score
585.3
Level
4

Item statistics	% OECD	% Ireland
Correct	40.3	40.8
Incorrect	42.3	46.5
Missing	17.4	12.7
Total	100	100

UNIT: STAIRCASE (2003)

Context: *Occupational*.

The diagram below illustrates a staircase with 14 steps and a total height of 252 cm:



Total height
252 cm

Total depth
400 cm

STAIRCASE QUESTION 1 (Item code: M547Q01)

Domain: *Space and shape.* **Item type**: *Short open constructed response.*

What is the height of each of the 14 steps? Height: ______ cm.

Key: *Full credit*: 18; *no credit*: Other responses, missing.
Process: *Reproduction.* Carry out a simple division; extract the relevant information from a single source.

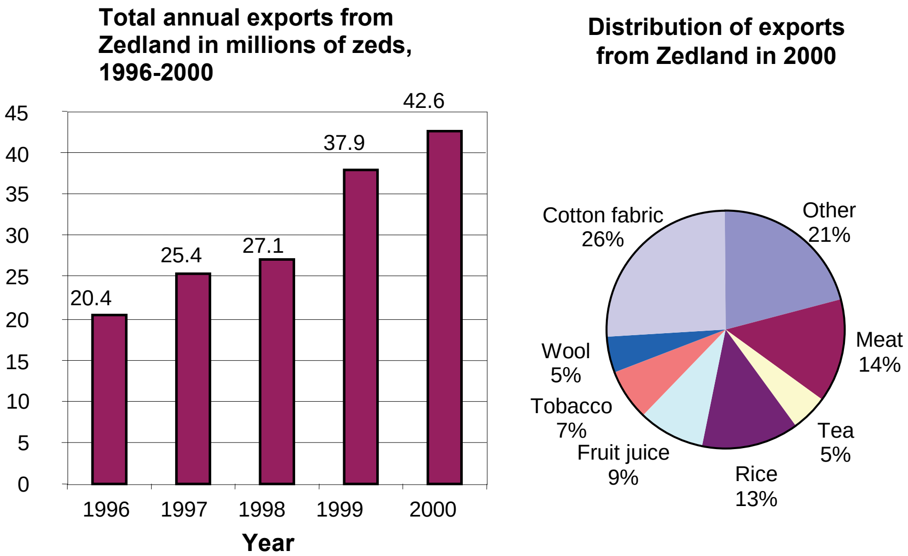
PISA Item Difficulty
Scale score
420.9
Level
2

Item statistics	% OECD	% Ireland
Correct	78.0	79.7
Incorrect	11.6	11.2
Missing	10.4	9.1
Total	100	100

UNIT: EXPORTS (2003)

Context: *Public*

The graphics below show information about exports from Zedland, a country that uses zeds as its currency.



EXPORTS QUESTION 1 (Item code: M438Q01)

Domain: *Uncertainty.* **Item type**: *Closed constructed response.*

What was the total value (in millions of zeds) of exports from Zedland in 1998?

Key: *Full credit*: 27.1 million zeds or 27 100 000 zeds or 27.1 (unit not required), accept also rounding to 27; *no credit*: Other responses.
Process: *Reproduction.* Follow the written instructions, decide which of the two graphs is relevant, and locate the correct information in that graph.

PISA Item Difficulty
Scale score
426.3
Level
2

Item statistics	% OECD	% Ireland
Correct	78.7	85.4
Incorrect	13.8	12.8
Missing	7.5	1.8
Total	100	100

EXPORTS QUESTION 2 (Item code: M438Q02)

Domain: *Uncertainty.* **Item type**: *Multiple choice.*

What was the value of fruit juice exported from Zedland in 2000?

A 1.8 million zeds
B 2.3 million zeds
C 2.4 million zeds
D 3.4 million zeds
E 3.8 million zeds

Key: *Full credit*: E (3.8 million zeds); *no credit*: Other responses, missing.
Process: *Connections.* Combine the information of two graphs, connect the numbers, and apply the appropriate basic mathematical routine.

PISA Item Difficulty
Scale score
565.0
Level
4

Item statistics	% OECD	% Ireland
Correct	48.3	50.8
Incorrect	44.8	46.2
Missing	6.9	3.0
Total	100	100

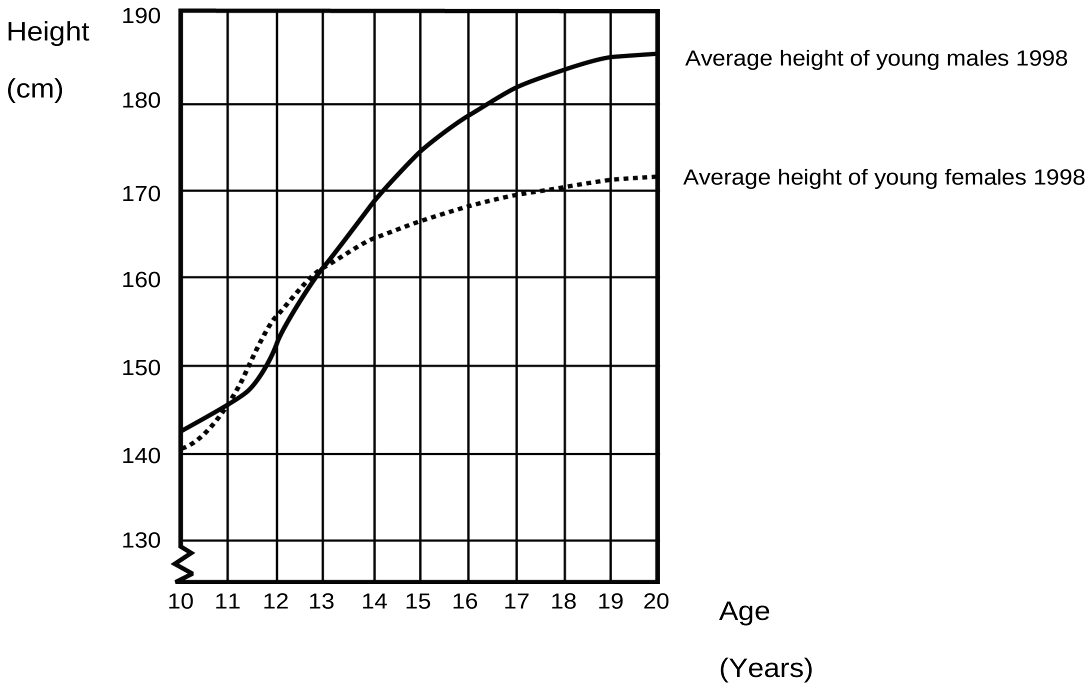
Response	% OECD	% Ireland
A	10.5	12.8
B	10.2	10.8
C	16.3	15.7
D	7.8	6.9
E*	48.3	50.8
Missing	6.9	3.03
Total	100	100

* Key.

UNIT: GROWING UP (2000)

Context: *Scientific.*

Young population grows taller



The average height of both young males and young females in the Netherlands in 1998 is represented in this graph.

GROWING UP QUESTION 1 (Item code: M150Q01)

Domain: *Change and relationships.* **Item type**: *Closed constructed response.*

Since 1980 the average height of 20-year-old females has increased by 2.3 cm, to 170.6 cm. What was the average height of a 20-year-old female in 1980? Answer: _____ cm.

Key: *Full credit*: 168.3 cm (unit already given); *no credit*: Other responses, missing.
Process: *Reproduction.* Extract the information from a single source, and ignore redundant information. Make use of a single representational mode, and employ a basic subtraction algorithm.

PISA Item Difficulty
Scale score
477.8
Level
2

Item statistics	% OECD	% Ireland
Correct	67.0	65.6
Incorrect	24.7	28.0
Missing	8.3	6.4
Total	100	100

GROWING UP QUESTION 2 (Item code: M150Q02)

Domain: *Change and relationships.* **Item type**: *Closed constructed response.*

According to this graph, on average, during which period in their life are females taller than males of the same age?

Key: *Full credit*: Gives the correct interval, from 11-13 years (using mathematical or daily-life language); *partial credit*: Other subsets of 11, 12, 13 years, not included in the full credit section; *no credit*: Other responses, missing.
Process: *Reproduction.* Interpret and decode standard representations of well known mathematical objects, compare the two graphs, and report the results.

PISA Item Difficulty
Scale score
419.3 (PC);525.3 (FC)
Level
between 1 and 2 (PC);
3 (FC)

Item statistics	% OECD	% Ireland
Fully correct	54.7	51.5
Partially correct	28.1	35.8
Incorrect	9.7	9.2
Missing	7.5	3.4
Total	100	100

GROWING UP QUESTION 3 (Item code: M150Q03)

Domain: *Change and relationships.* **Item type**: *Open constructed response.*

Explain how the graph shows that on average the growth rate for girls slows down after 12 years of age.

Key: *Full credit*: The response should refer to the "change" of the gradient of the graph for female (explicitly or implicitly, in mathematical language or using daily-life language), or the student should mention that the female graph becomes less steep, as well as the fact that the graph falls below the male graph; *no credit*: Student indicates that female height drops below male height, but does not mention the steepness of the female graph or makes a comparison of the female growth rate before and after 12 years, other responses, missing.
Process: *Connections.* Combine 'growing' and 'slowing down', link different ideas and information, and communicate the results.

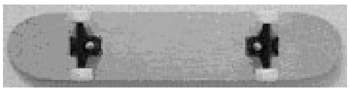
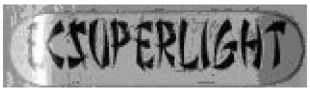
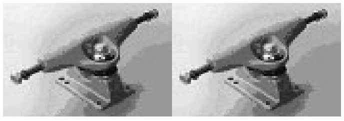
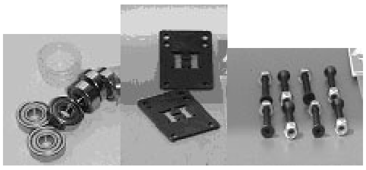
PISA Item Difficulty
Scale score
573.6
Level
4

Item statistics	% OECD	% Ireland
Correct	44.8	56.5
Incorrect	34.1	32.8
Missing	21.1	10.7
Total	100	100

UNIT: SKATEBOARD (2003)

Context: *Personal.*

Eric is a great skateboard fan. He visits a shop called SKATERS to check some prices. At this shop you can buy a complete board. Or you can buy a deck, a set of 4 wheels, a set of 2 trucks and a set of hardware, and assemble your own board. The prices for the shop's products are:

Product	Price in zeds	
Complete skateboard	82 or 84	
Deck	40, 60 or 65	
One set of 4 Wheels	14 or 36	
One set of 2 Trucks	16	
One set of hardware (bearings, rubber pads, bolts and nuts)	10 or 20	

SKATEBOARD QUESTION 1 (Item code: M520Q01b)

Domain: *Quantity.* **Item type**: *Short constructed response.*

Eric wants to assemble his own skateboard. What is the minimum price and the maximum price in this shop for self-assembled skateboards?

(a) Minimum price: _______________ zeds.
(b) Maximum price: _______________ zeds.

Key: *Full credit*: Both the minimum (80) and the maximum (137) are correct; *partial credit*: Only the minimum (80) is correct, or only the maximum (137) is correct; *no credit*: Other responses, missing.
Process: *Reproduction.* Find a simple strategy to come up with the maximum and minimum, use of a routine addition procedure, use of a simple table.

PISA Item Difficulty
Scale score
463.7 (PC); 496.5 (FC)
Level
2 (PC); 3 (FC)

Item statistics	% OECD	% Ireland
Fully correct	66.7	69.0
Partialy correct	10.6	8.2
Incorrect	18.0	20.8
Missing	4.7	2.0
Total	100.0	100.0

SKATEBOARD QUESTION 2 (Item code: M520Q02)

Domain: *Quantity.* **Item type**: *Multiple choice.*

The shop offers three different decks, two different sets of wheels and two different sets of hardware. There is only one choice for a set of trucks.
How many different skateboards can Eric construct?

A 6
B 8
C 10
D 12

Key: *Full credit*: D; *no credit*: Other responses, missing.
Process: *Reproduction.* Interpret a text in combination with a table correctly; apply a simple enumeration algorithm accurately.

PISA Item Difficulty
Scale score
569.7
Level
4

Item statistics	% OECD	% Ireland
Correct	45.5	30.2
Incorrect	50.0	66.9
Missing	4.5	2.9
Total	100	100

Response	% OECD	% Ireland
A	25.4	33.2
B	18.3	27.7
C	6.3	6.0
D*	45.5	30.2
Missing	4.5	2.9
Total	100	100

*Key.

SKATEBOARD QUESTION 3 (Item code: M520Q03)

Domain: *Quantity.* **Item type**: *Short constructed response.*

Eric has 120 zeds to spend and wants to buy the most expensive skateboard he can afford. How much money can Eric afford to spend on each of the 4 parts? Put your answer in the table below.

Part	**Amount (zeds)**
Deck	*65 zeds*
Wheels	*14 zeds*
Trucks	*16 zeds*
Hardware	*20 zeds*

Key: *Full credit*: See in table above; *no credit*: Other responses, missing.
Process: *Connections.* Relate text based information to a table representation, apply a non-standard strategy, and carry out routine calculations.

PISA Item Difficulty
Scale score
554.1
Level
4

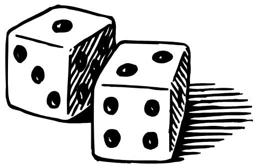
Item statistics	**% OECD**	**% Ireland**
Fully correct	49.8	50.3
Incorrect	44.7	47.9
Missing	5.5	1.8
Total	100	100

Number of correct responses	**% OECD**	**% Ireland**
0	5.1	3.7
1	5.6	5.9
2	17.3	19.4
3	16.7	18.9
4	49.8	50.3
Missing	5.5	1.8
Total	100	100

UNIT: NUMBER CUBES (2003)

Context: *Personal.*

On the right, there is a picture of two dice.



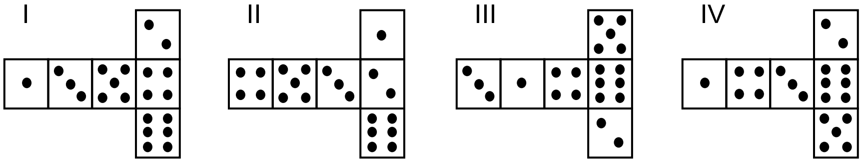
Dice are special number cubes for which the following rule applies:
The total number of dots on two opposite faces is always seven. You can make a simple number cube by cutting, folding and gluing cardboard. This can be done in many ways.

NUMBER CUBES QUESTION 2 (Item code: M555Q02)

Domain: *Space and shape.* **Item type**: *Complex multiple choice.*

In the figure below you can see four cuttings that can be used to make cubes, with dots on the sides.

Which of the following shapes can be folded together to form a cube that obeys the rule that the sum of opposite faces is 7? For each shape, circle either "Yes" or "No" in the table below.



Shape	Obeys the rule that the sum of opposite faces is 7?
I	Yes / No
II	Yes / No
III	Yes / No
IV	Yes / No

Key: *Full credit*: No, yes, yes, and no, in that order; *no credit*: Other responses, missing.
Process: *Connections.* Encode and interpret 2-dimensional objects, interpret the connected 3-dimensional object, and check certain basic computational relations.

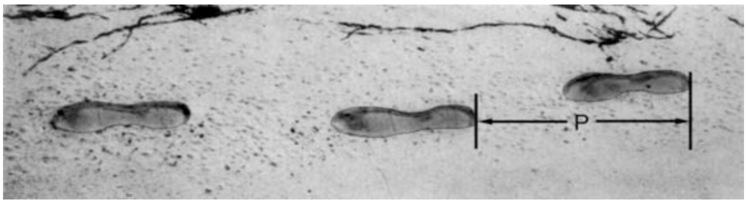
PISA Item Difficulty
Scale score
503.5
Level
3

Item statistics	% OECD	% Ireland
Correct	63.0	57.4
Incorrect	34.7	40.9
Missing	2.3	1.7
Total	100	100

Number of correct responses	% OECD	% Ireland
0	2.7	3.1
1	7.2	8.9
2	8.9	8.2
3	16.0	20.7
4	63.0	57.4
Missing	2.3	1.7
Total	100	100

UNIT: WALKING (2000)

Context: *Personal.*



The picture shows the footprints of a man walking. The pace length *P* is the distance between the rear of two consecutive footprints. For men, the formula, n/P = 140, gives an approximate relationship between *n* and *P* where *n* = number of steps per minute and *P* = pace length in metres.

WALKING QUESTION 1 (Item code: M124Q01)

Domain: *Change and relationships.* **Item type**: *Open constructed response.*

If the formula applies to Mark's walking and Mark takes 70 steps per minute, what is Mark's pace length? Show your work.

Key: *Full credit*: 0.5 m or 50 cm, ½ (unit not required); *partial credit*: 70/ *p* = 140, 70 = 140 *p*, *p* = 0.5. 70/140; *no credit*: Other responses, missing.
Process: *Reproduction.* Reflect on and realise the embedded mathematics, solve the problem successfully through substitution in a simple formula, and carry out a routine procedure.

PISA Item Difficulty
Scale score
611.0
Level
5

Item statistics	% OECD	% Ireland
Fully Correct	36.3	22.9
Partially Correct	21.8	34.7
Incorrect	20.9	28.1
Missing	21.0	14.3
Total	100	100

WALKING QUESTION 3 (Item code: M124Q03)

Domain: *Change and relationships.* **Item type**: *Open constructed response.*

Bernard knows his pace length is 0.80 metres. The formula applies to Bernard's walking. Calculate Bernard's walking speed in metres per minute and in kilometres per hour. Show your working out.

Key: *Full credit*: Correct answers (unit not required) for both metres/minute and km/hour: n = 140 x .80 = 112. Per minute he walks 112 x .80 metres = 89.6 metres, or as long as both correct answers are given (89.6 and 5.4), whether working out is shown or not. Errors due to rounding are acceptable; *partial credit (2-point)*: Student fails to multiply by 0.80 to convert from steps per minute to metres per minute, or if the speed in metres per minute correct (89.6 metres per minute) but conversion to kilometres per hour incorrect or missing, or correct method (explicitly shown) with minor calculation error(s) with no answers correct, or only 5.4 km/hr is given, but not 89.6 metres/minute (intermediate calculations not shown); *partial credit (1-point)*: *n* = 140 x .80 = 112. No further working out is shown or incorrect working out from this point; *no credit*: Other responses, missing.
Process: *Connections.* Substitute in an algebraic expression, do a sequence of different but connected calculations that need proper understanding of transforming formulas and units of measures.

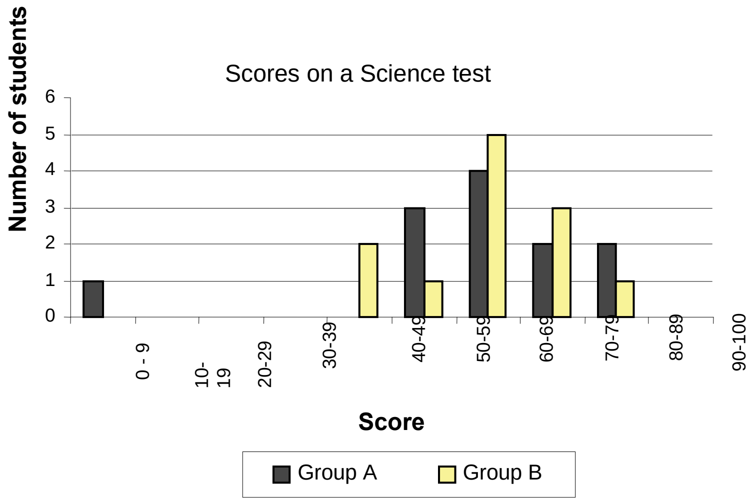
PISA Item Difficulty
Scale score
604.7 (PC 1-point);
666.3 (PC 2-point);
722.3 (FC)
Level
4 (PC 1-point);
5 (PC 2-point); 6 (FC)

Item statistics	% OECD	% Ireland
Fully correct	8.0	3.7
Partially correct (2-point)	9.0	4.8
Partially correct (1-point)	19.9	20.4
Incorrect	24.4	39.1
Missing	38.7	31.9
Total	100	100

UNIT: TEST SCORES (2003)

Context: *Educational.*

The diagram below shows the results on a Science test for two groups, labelled as Group A and Group B.
The mean score for Group A is 62.0 and the mean for Group B is 64.5. Students pass this test when their score is 50 or above.



TEST SCORES QUESTION 1 (Item code: M513Q01)

Domain: *Uncertainty.* **Item type**: *Open constructed.*

Looking at the diagram, the teacher claims that Group B did better than Group A in this test.
The students in Group A don't agree with their teacher. They try to convince the teacher that Group B may not necessarily have done better.
Give one mathematical argument, using the graph that the students in Group A could use.

Key: *Full credit*: One valid argument is given. Valid arguments could relate to the number of students passing, the disproportionate influence of the outlier, or the number of students with scores in the highest level; *no credit*: Other responses, including responses with no mathematical reasons, or wrong mathematical reasons, or responses that simply describe differences but are not valid arguments, missing.
Process: *Connections.* Apply statistical knowledge in a problem situation where the mathematical representation is partially apparent, interpret and analyse given information, and communicate reasons and arguments.

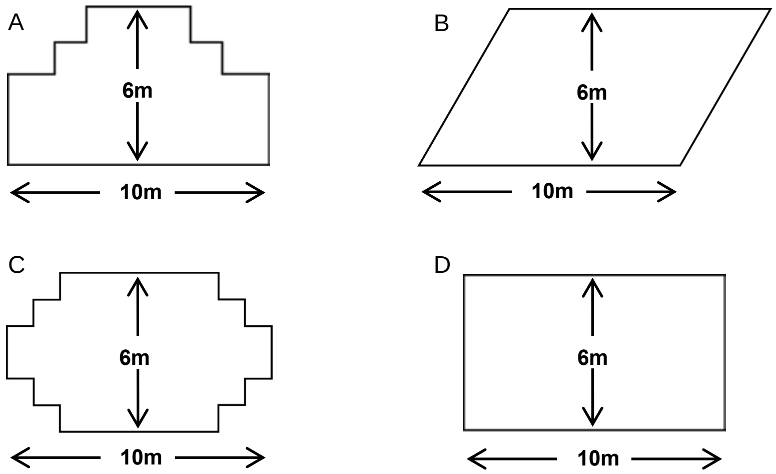
PISA Item Difficulty
Scale score
619.5
Level
5

Item statistics	% OECD	% Ireland
Correct	32.2	40.8
Incorrect	32.8	38.6
Missing	35.0	20.6
Total	100	100

UNIT: CARPENTER (2000)

Context: *Educational.*

A carpenter has 32 metres of timber and wants to make a border around a vegetable patch. He is considering the following designs for the vegetable patch.



CARPENTER QUESTION 1 (Item code: M266Q01)

Domain: *Space and shape.* **Item type**: *Complex multiple choice.*

Circle either "Yes" or "No" for each design to indicate whether the vegetable patch can be made with 32 metres of timber.

Vegetable patch design	Using this design, can the vegetable patch be made with 32
Design A	Yes / No
Design B	Yes / No
Design C	Yes / No
Design D	Yes / No

Key: *Full credit*: Four correct (yes, no, yes, yes, in that order); *partial credit*: Three correct; *no credit*: Two or fewer correct; missing.
Process: *Connections.* Use geometrical insight and argumentation skills, and possibly some technical geometrical knowledge.

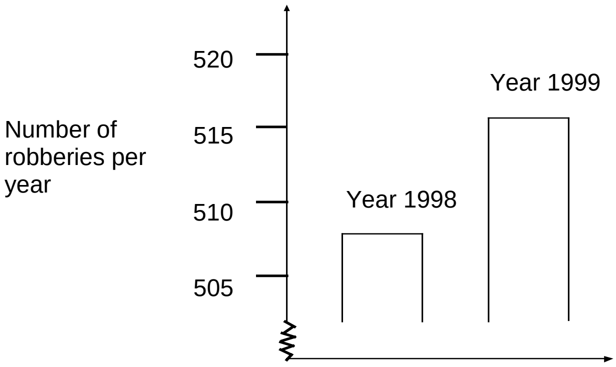
PISA Item Difficulty
Scale score
687.3
Level
6

Item statistics	% OECD	% Ireland
Fully correct	20.0	13.0
Partially correct	30.8	30.9
Incorrect	46.8	54.6
Missing	2.5	1.6
Total	100	100

UNIT: ROBBERIES (2000)

Context: *Public.*

A TV reporter showed this graph to the viewers and said:
"The graph shows that there is a huge increase in the number of robberies from 1998 to 1999."



ROBBERIES QUESTION 1 (Item code: M179Q01)

Domain: *Uncertainty.* **Item type**: *Open constructed response.*

Do you consider the reporter's statement to be a reasonable interpretation of the graph? Give an explanation to support your answer.

Key: *Full credit*: "No, not reasonable". Focuses on the fact that only a small part of the graph is shown; *partial credit*: "No, not reasonable", but explanation lacks detail, or "No, not reasonable", with correct method but with minor computational errors; *no credit*: No, with no, insufficient or incorrect explanation, yes, other responses, missing.
Process: *Connections.* Focus on an increase given by an exact number of robberies in absolute and relative terms; argumentation based on interpretation of data.

PISA Item Difficulty
Scale score
576.7 (PC); 694.3 (FC)
Level
4 (PC); 6 (FC)

Item statistics	% OECD	% Ireland
Fully correct	15.4	13.3
Partially correct	28.1	36.7
Incorrect	41.5	38.1
Missing	15.0	11.9
Total	100	100

PROBLEM SOLVING

UNIT: CINEMA OUTING (2003)

Problem type: *Decision making.*

This problem is about finding a suitable time and date to go to the cinema.
Jason, a 15-year-old, wants to organise a cinema outing with two of his friends, who are of the same age, during the one-week school holiday. The holidays begin on Saturday, 24th March and end on Sunday, 1st April. Jason asks his friends for suitable dates and times for the outing. The following information is what he received.
Fred: "I've to stay home on Monday and Wednesday afternoons for music practice between 2:30 and 3:30."
Simon: "I've to visit my grandmother on Sundays, so it can't be Sundays. I have seen Pokamin and don't want to see it again."
Jason's parents insist that he only goes to films suitable for his age and does not walk home. They will fetch the boys home at any time up to 10 p.m.
Jason checks the film times for that week. This is the information that he finds.

TIVOLI CINEMA

Advance Booking Number: 1850 2003545
24 hour phone number: 1850 2020200
Bargain Day Tuesdays: All films €5

Films showing from Fri 23rd March for two weeks:

Children in the Net		**Pokamin**	
113 mins 14:00 (Mon-Fri only) 21:35 (Sat/Sun only)	Suitable only for persons of 12 years and over	105 mins 13:40 (Daily) 16:35 (Daily)	Parental Guidance. General viewing, but some scenes may be unsuitable for young children
Monsters from the Deep		**Enigma**	
164 mins 19:55 (Fri/Sat only)	Suitable only for persons of 18 years and over	144 mins 15:00 (Mon-Fri only) 18:00 (Sat/Sun only)	Suitable only for persons of 12 years and over
Carnivore		**King of the Wild**	
148 mins 18:30 (Daily)	Suitable only for persons of 18 years and over	117 mins 14:35 (Mon-Fri only) 18:50 (Sat/Sun only)	Suitable for persons of all ages

CINEMA OUTING QUESTION 1 (Item code: X601Q01)

Item type: *Multiple choice.*

Taking into account the information Jason found on the films, and the information he got from his friends, which of the six films should Jason and the boys consider watching?
Circle "Yes" or "No" for each film.

Film	Should the three boys consider watching the film?
Children in the Net	Yes / No
Monsters from the Deep	Yes / No
Carnivore	Yes / No
Pokamin	Yes / No
Enigma	Yes / No
King of the Wild	Yes / No

Key: *Full credit:* Yes, no, no, no, yes, yes, in that order; *partial credit:* One incorrect answer; *no credit:* Other responses, missing.
Process: Not classified by process.

PISA Item Difficulty
Scale score
441.9 (PC); 521.9 (FC)
Level
1 (PC); 2 (FC)

Item statistics	% OECD	% Ireland
Fully correct	55.5	64.3
Partially correct	23.4	18.5
Incorrect	19.1	16.0
Missing	2.0	1.2
Total	100	100

Note. Fully correct and partially correct are combined in the PISA International Report.

CINEMA OUTING QUESTION 2 (Item code: X601Q02)

Item type: *Multiple choice.*

If the three boys decided on going to "Children in the Net", which one of the following dates is suitable for them?

A Monday, 26th March
B Wednesday, 28th March
C Friday, 30th March
D Saturday, 31st March
E Sunday, 1st April

Key: *Full credit:* C; *no credit:* Other responses, missing.
Process: Not classified by process.

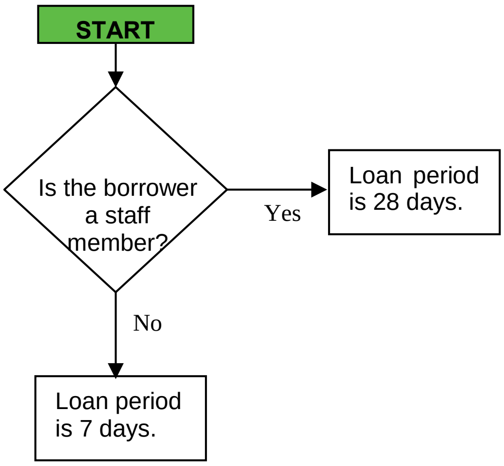
PISA Item Difficulty
Scale score
468.3
Level
1

Item statistics	% OECD	% Ireland
Correct	68.1	77.8
Incorrect	19.1	15.3
Missing	12.8	6.9
Total	100	100

UNIT: LIBRARY SYSTEM (2003)

Problem type: *System analysis and design.*

The **Moatstown Community School** library has a simple system for lending books: for staff members the loan period is 28 days, and for students the loan period is 7 days. The following is a flow chart showing this simple system:



The **Dunbeg Secondary School** library has a similar, but more complicated, lending system:

- All publications classified as "Reserved" have a loan period of 2 days.
- For books (not including magazines) that are **not** on the reserved list, the loan period is 28 days for staff, and 14 days for students.
- For magazines that are **not** on the reserved list, the loan period is 7 days for everyone.
- Persons with any overdue items are not allowed to borrow anything.

LIBRARY SYSTEM QUESTION 1 (Item code: X402Q01)

Item type: *Closed constructed response.*

You are a student at **Dunbeg Secondary School**, and you do not have any overdue items from the library. You want to borrow a book that is **not** on the reserved list. How long can you borrow the book for? Answer: _______ days.

Key: *Full credit:* 14 days; *no credit:* Other responses, missing.
Process: Not classified by process.

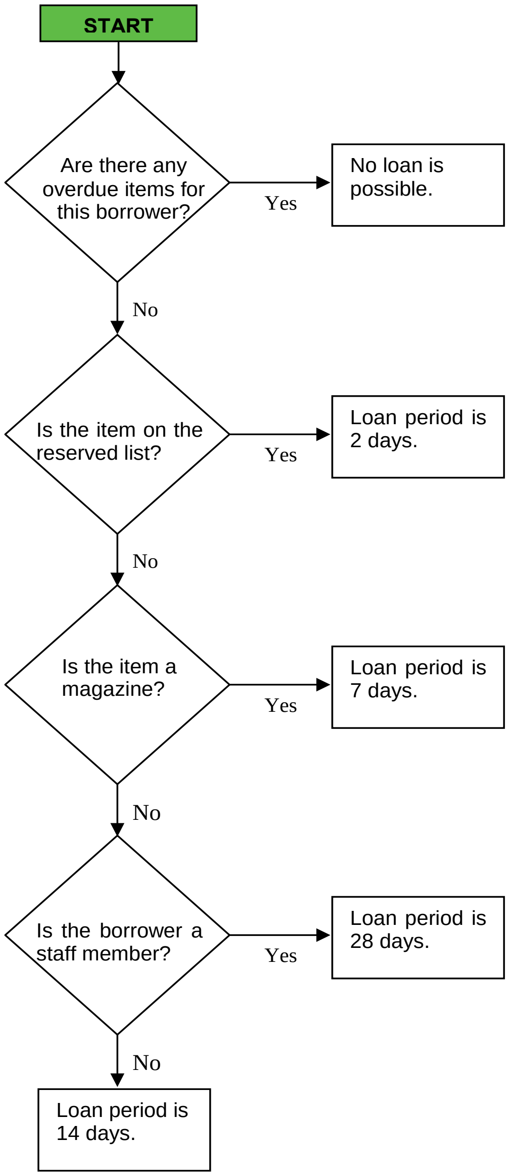
PISA Item Difficulty
Scale score
436.8
Level
1

Item statistics	% OECD	% Ireland
Correct	74.8	86.8
Incorrect	19.5	10.9
Missing	5.7	2.3
Total	100	100

LIBRARY SYSYEM QUESTION 2 (Item code: X402Q02)

Item type: *Open constructed response.*

Develop a flow chart for the **Dunbeg Secondary School** Library system so that an automated checking system can be designed to deal with book and magazine loans at the library. Your checking system should be as efficient as possible (i.e. it should have the least number of checking steps). Note that each checking step should have only **two** outcomes and the outcomes should be labelled appropriately (e.g. "Yes" and "No").



Key: *Full credit:* The most efficient system is the 4-step check system as above; equivalent statements can be accepted:

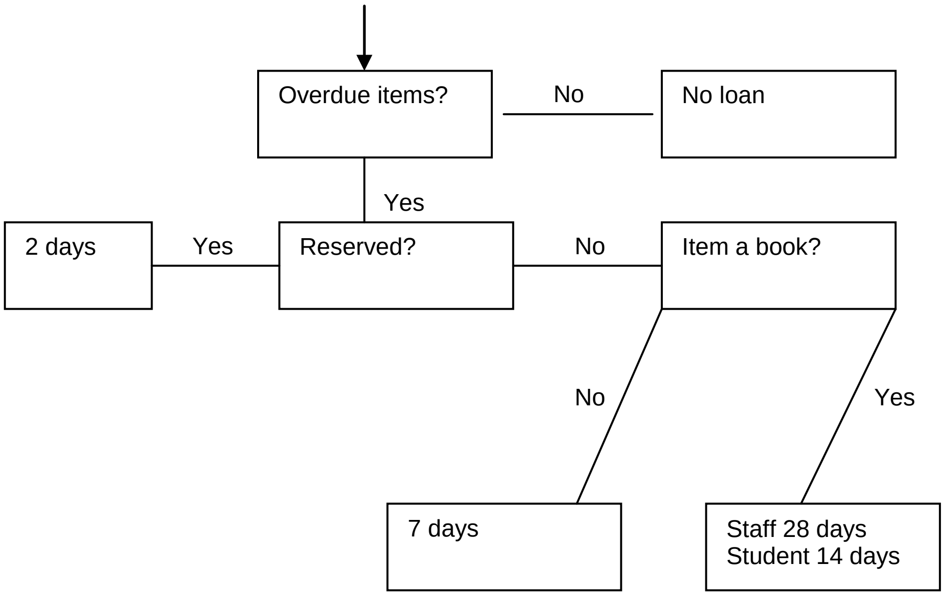
Partial credit (1):

The diagram is correct except that the first three check steps are out of order in one (but not both) of the following two ways: The checks for "reserved list" and "magazine" are interchanged. The checks for "overdue items" and "reserved list" are interchanged.

The check for "overdue items" is written as a statement outside the flow chart.

The other three check steps are in the right sequence, but with a "minor error".

The check for "overdue items" is missing, but the other three check steps are completely correct and in the right sequence.
Partial credit (2):
The four check steps are in the right sequence, but there is a "minor error". For example: One loan period is incorrect; one loan period is missing; one or more Yes/No missing;
one Yes/No incorrectly labelled. For example:



Or: The check for "overdue items" is written as a statement outside the flow chart, but the other three check steps are completely correct and in the right sequence.
Two check steps are out of order, resulting in 5 steps, as one extra check step is required.
The system is still "complete", but less efficient. By "complete" we mean that the checking system will produce the correct loan periods in all cases.
No credit: The system is "complete", but has more than 5 check steps; other responses, missing.
Process: Not classified by process.

PISA Item Difficulty
Scale score
658.1 (PC1); 677.8 (PC2);
693.0 (FC)
Level
3 (PC1, PC2, FC)

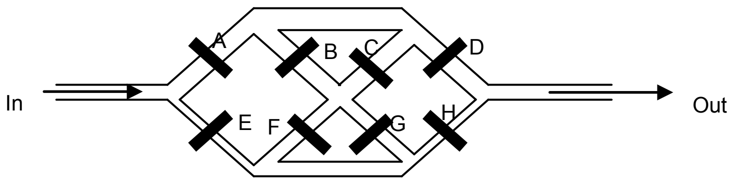
Item statistics	% OECD	% Ireland
Fully correct	9.8	3.3
Part. correct (2)	3.5	3.4
Part. correct (1)	6.8	4.5
Incorrect	56.8	74.2
Missing	23.2	14.7
Total	100	100

UNIT: IRRIGATION (2003)

Problem type: *Trouble shooting.*

Below is a diagram of a system of irrigation channels for watering sections of crops. The gates A to H can be opened and closed to let the water go where it is needed. When a gate is closed no water can pass through it. This is a problem about finding a gate which is stuck closed, preventing water from flowing through the system of channels.

Figure 1: A system of irrigation channels



Michael notices that the water is not always going where it is supposed to.
He thinks that one of the gates is stuck closed, so that when it is switched to "open", it does not open.

IRRIGATION QUESTION 1 (Item code: X603Q01)

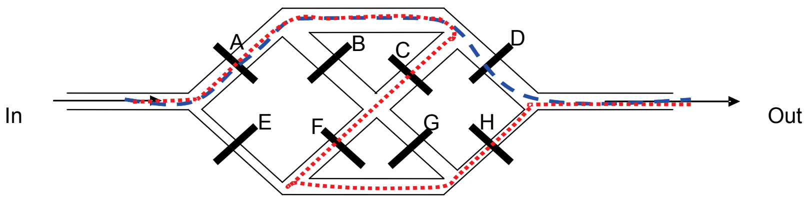
Item type: *Open constructed response.*

Michael uses the settings given in Table 1 to test the gates.

Table 1: Gate Settings

A	B	C	D	E	F	G	H
Open	Closed	Open	Open	Closed	Open	Closed	Open

With the gate settings as given in Table 1, on the diagram below draw all the possible paths for the flow of water. Assume that all gates are working according to the settings.



Key: *Full credit:* Flow paths as shown above: Ignore any indications of the directions of flow; the response could be shown in the diagram provided, or in figure 1, or in words, or with arrows; *no credit:* Other responses, missing.
Process: Not classified by process.

PISA Item Difficulty
Scale score
497.2
Level
1

Item statistics	% OECD	% Ireland
Correct	47.1	45.9
Incorrect	50.2	52.3
Missing	2.7	1.8
Total	100	100

IRRIGATION QUESTION 2 (Item code: X603Q02)

Item type: *Multiple choice.*

Michael finds that, when the gates are set as shown in Table 1, no water flows through, indicating that at least one of the gates set to "open" is stuck closed.
Decide for each problem case below whether the water will flow through all the way. Circle "Yes" or "No" in each case.

Problem Case	Will water flow through all the way?
Gate **A** is stuck closed. All other gates are working properly as set in Table 1.	Yes/No
Gate **D** is stuck closed. All other gates are working properly as set in Table 1.	Yes/No
Gate **F** is stuck closed. All other gates are working properly as set in Table 1.	Yes/No

Key: *Full credit:* No, yes, yes, in that order; *no credit:* Other responses, missing.
Process: Not classified by process.

PISA Item Difficulty
Scale score
543.2
Level
2

Item statistics	% OECD	% Ireland
Correct	36.1	35.6
Incorrect	25.1	32.5
Missing	38.8	31.9
Total	100	100

IRRIGATION QUESTION 3 (Item code: X603Q03)

Item type: *Open constructed response.*

Michael wants to be able to test whether **gate D** is stuck closed.
In the following table, show settings for the gates to test whether **gate D** is stuck closed when it is set to "open".

Settings for gates (each one "open" or "closed"):

A	B	C	D	E	F	G	H

Key: *Full credit:* A and E are not both closed. D must be open. H can only be open if water cannot get to it (e.g., other gates are closed preventing water from reaching H). Otherwise H must be closed (H closed, all other gates open); *no credit:* Other responses, missing.
Process: Not classified by process.

PISA Difficulty
Scale score
531.3
Level
2

Item statistics	% OECD	% Ireland
Correct	54.4	55.8
Incorrect	32.9	35.1
Missing	12.7	9.1
Total	100	100

Appendix B

B.1 Description of Student and School Level Variables

STUDENT LEVEL
Background Variables

Name	Description	Data Type	Discrete Value/Descriptive	% Missing
Gender				
	Whether the student is male or female	Categorical (2)	Male Female	0.0
Nationality				
	Native status of student classified into three categories	Categorical (3)	Native First generation Non-native	1.6
Membership of Traveller Community				
	Is student member of the Traveller community	Categorical (2)	Yes No	6.1
Socioeconomic Status				
Parents' SES				
	Parents' SES, based on the highest occupation of mother or father, coded and scaled according to ISCO, then ISEI.	Continuous	r=16 to 91; M=48.34; SD=15.87.	4.3
	Divided into three categories	Ordered categorical (3)	Low Medium High	
Parents' Educational Attainment				
	Parents' highest level of educational attainment (ISCED)	Ordered categorical (5)	None/primary Lower secondary Post-secondary/ non-tertiary Tertiary cert/diploma Tertiary degree	2.5
Economic, Social and Cultural Status				
	Index derived from three variables: highest parent ISEI; highest parent educational attainment; and cultural possessions (e.g., number of books in the home).	Continuous	r=-4.61to 2.22; M=-.081; SD=0.89.	1.5
	Composite score, with three categories.	Ordered Categorical (3)	Low Medium High	
Family Structure				
Lone-Parent Status[60]				
	Whether student lives in a single-parent household	Categorical (2)	Yes No	3.6

[60] Derived from household composition variable originally categorised as lone-parent, nuclear and mixed families. Around 2% of students indicated that they were living in 'other' household compositions. Mean achievement of these students did not differ from those with missing data. Therefore, other was combined with missing.

Number of Siblings				
	Number of brothers and sisters	Continuous	r=0 to 17; M=2.66; SD=2.08.	2.2
	Derived from number of brothers and sisters	Ordered categorical (4)	One; Two; Three; Four or more	
Home Educational Resources				
Number of Educational Resources				
	Number of educational resources in the home, including a desk and a quiet place for study, and books to help with school work.	Ordered categorical (3)	Low (None or One) Medium (Two) High (Three)	1.2
Books in the Home				
	Number of books in the home	Ordered categorical (6)	0-10; 11-25; 26-100; 101-200; 201-500; 500+	2.1

Student Out of School Activities

Name	*Description*	*Data Type*	*Discrete Value/Descriptive*	*% Missing*
Homework Practices				
Total Time Spent on Homework/Study				
	Number of hours homework and study per week	Continuous	r=0 to 30; M=7.73; SD=5.69.	8.5
Total Time Spent on Mathematics Homework/Study				
	Number of hours spent on Mathematics homework and study per week	Continuous	r=0 to 24; M=2.84; SD=2.38.	9.4
Leisure Reading				
Reading Fiction				
	Frequency of reading fiction	Categorical (5)	Hardly ever/never Few times a yr Once a mth Once a wk Several times wk	5.6
Reading Emails/Webpages				
	Frequency of reading emails/webpages	Categorical (5)	Hardly ever/never Few times a yr Once a mth Once a wk Several times wk	6.3

Student Academic Characteristics and Behaviour

Name	*Description*	*Data Type*	*Discrete Value/Descriptive*	*% Missing*
Current Grade Level				
	Current grade level of student	Ordered categorical (4)	Mode: 3rd year 2nd; 3rd; 4th/Transition Year; 5th	0.0
Absence from School				
	Frequency of missing school in past two weeks	Ordered categorical (4)	None;1-2; 3-4; 5+	3.3

Risk of Early School Leaving				
	Whether student intends to leave school before completion of Leaving Certificate Examination.	Categorical (2)	Yes No	1.1
Calculator usage in PISA				
	Whether student used calculator during PISA assessment	Categorical (2)	Yes No	12.3
Study of Science				
	Whether or not student was studying, or had studied science for the Junior Cert Examination.	Categorical (2)	Yes No	1.7
Syllabus Level				
	Syllabus level that students had taken, or intended to take, for the Junior Certificate Examination	Categorical (4)	Higher, Ordinary, Foundation, do not study (science)	0.6 - 1.7 (Depending on JCE subject)

Student Beliefs about Mathematics

Name	*Description*	*Data Type*	*Discrete Value/Descriptive*	*% Missing*
Self-Efficacy in Mathematics				
	Eight items measuring students' confidence with specific mathematical tasks were used to create a composite measure of mathematical efficacy	Continuous composite	r= -3.89 to 2.53; M = -0.03; SD = 0.94.	1.6
		Ordered categorical (3)	Low Medium High	
Anxiety towards Mathematics				
	Five items measuring general concerns about achievements in mathematics were used to create a composite measure of mathematical anxiety	Continuous composite	r= -2.48 to 2.70, M = 0.07, SD = 0.93	1.8
		Ordered categorical (3)	Low Medium High	

SCHOOL LEVEL

School Structure

Name	*Description*	*Data Type*	*Discrete Value/Parameter*	*% Missing*
Stratum (Size)				
	Sampling stratum based on enrolment of 15-year-olds	Ordered categorical (3)	Small Medium Large	0.0
Sector				
	Classification of schools based on sector	Categorical (3)	Secondary Community/ Comprehensive Vocational	0.0

Disadvantaged Status				
	Using the DES's database, schools were categorised based on their inclusion in the Disadvantaged Area Schools Scheme (DAS)	Categorical (2)	Disadvantaged Not Disadvantaged	0.0
Economical, Social and Cultural Status (ESCS)				
	School average for student ESCS measure (see above; student background variables)	Continuous composite	r=-1.39 to 1.28 M=-0.79 SD=0.43	0.0
	Divided into three categories	Ordered categorical (3)	Low Medium High	
Gender Composition				
	Using the DES's post-primary schools database, schools were categorised according to the proportion of female 15-year-olds	Ordered categorical (4)	0%; 0.1-45%; 45.1%-99.9%; 100%	0.0
Fee Waiver Entitlement				
	Weighted percentage of students entitled to fee waiver for the JCE	Continuous	r=0 to 73.68; M=25.84; SD=15.21	0.0

School Climate and Resources

Name	*Description*	*Data Type*	*Discrete Value/Descriptive*	*% Missing*
Disciplinary Climate				
	Based on student responses regarding the frequency with which things happen in their mathematics classroom (e.g., noise, attention, disruption, obedience etc.), a composite of disciplinary climate in was created. This was aggregated to the school-level, and disaggregated at the student-level.	Continuous composite	r=-2.74 to 2.35; M=0.27;SD=1.15 (student-level data, based on school aggregate level)	0.0
	Divided into three categories.	Ordered categorical (3)	Low Medium High	
Ratio of Computers to Students				
	Variable computed by dividing the number of computers available to 15-year-olds in a school by the number of 15-year-olds in the school.	Ratio	r=0.002 to 0.356; M=0.112; SD=0.068	16.1
	Divided into three categories	Ordered categorical (3)	Low Medium High	

Instructional Time				
	Total minutes of instruction time per week	Continuous	r=315 to 3160 M=1643 SD=300.8	19.7
		Ordered categorical (3)	Low Medium High	

B.2 Mapping the Irish Education System onto the International Standard Classification of Education (ISCED)

In PISA, education programmes are classified according to the *International Standard Classification of Education* (ISCED) which distinguishes programmes according to level, orientation and designation (See OECD, 1999). Orientation is classified as mainly academic ('a'), mainly technical/vocational ('b'), or 'terminal' qualifications for direct access to the labour market ('c'). The following levels are relevant in interpreting parental education of Irish students:

- ISCED Level 1: Primary level of education (first class to sixth class or first grade to sixth grade);
- ISCED Level 2: Lower secondary level of education (Junior Cycle; first year to third year or grade 7 to grade 9. In Ireland, all students take the same 'a', or academic-oriented Junior Cycle; there are no 'b' or 'c' programmes, see below for an explanation of those programmes);
- ISCED Level 3: Upper secondary level of education (Senior Cycle; fourth year to sixth year or grade 10 to grade 12. In Ireland, the traditional Leaving Certificate and the Leaving Certificate Vocational Programme are both classified as 'a', or academic in orientation, while the Transition Year programme and Leaving Certificate Applied are classified as 'c' in orientation, though it should be acknowledged that many students taking transition year progress to an ISCED 3A programme, and then to tertiary education; there are no 'b' programmes);
- ISCED Level 4: Qualification obtained in programmes that straddle the boundary between upper-secondary and post-secondary education. They are typically not significantly more advanced than programmes at Level 3 and have a full-time equivalent duration of between 6 months and 2 years (e.g., short secretarial/technical training course);
- ISCED Level 5A: Qualification obtained from a tertiary study programme with a strong theoretical foundation, typically with a minimum duration of three years' full time equivalent, providing entry into profession with high skills requirements or an advanced research programme, and involving completion of a research project or thesis (e.g., BA, BSc);
- ISCED Level 5B: Qualification obtained in tertiary programmes that are generally more practical/technical/occupationally specific and typically shorter than ISCED 5A programmes. Typically, these programmes have a minimum of two years' full-time equivalent duration and prepare students to enter a particular occupation (e.g., IT Cert. in bar management, IT cert. in hospitality management); and
- ISCED Level 6: Advanced research qualification typically obtained after a second stage of tertiary education (e.g., MSc, PhD).

In the analysis of parents' education reported in Chapter 4, parental education levels are classified as none, ISCED 1, ISCED 2, ISCED 3a, ISCED 4, ISCED 5b, and ISCED 5a/6. (ISCED 3c programme were not available to the parents of students in PISA 2003).

There are three sub-divisions at each of Levels 2 and 3:

- Level 2A comprises programmes designed to prepare students for direct access to Level 3 in a sequence which would ultimately lead to tertiary education. The Junior Certificate Programme is categorised as this level;
- Level 2B comprises programmes designed to prepare students for direct access to programmes at Level 3C. There are no Irish programmes classified as Level 2B;
- Level 2C programmes are designed primarily for direct access to the labour market at the end of Level 2 (lower secondary level). There are no Irish programmes classified as Level 2C;
- Level 3A comprises programmes designed to provide direct access to primarily academically oriented tertiary programmes. It corresponds to the Leaving Certificate and Leaving Certificate Vocational Programmes;
- Level 3B programmes are designed to provide direct access to technically- or vocationally-oriented tertiary programmes. There are no Irish programmes classified as Level 3B; and
- Level 3C programmes lead directly to labour market or to post-primary non-tertiary programmes. The Transition Year and Leaving Certificate Applied Programmes are classified as Level 3C, though it must be acknowledged that many students who complete the Transition Year proceed to tertiary education.

B.3 Procedure for Computing Estimates and Standard Errors

As computer packages such as Wesvar do not provide standard errors associated with some parameters, such as the percentage of students at each proficiency level, it was necessary to compute such standard errors using the following procedure.

1. Separate percentages for groups of students (e.g., females, males) were computed in Wesvar, for each plausible value. Hence, five percentages were estimated (one associated with each plausible value). Each set (P1 to P5) was then averaged to provide a mean parameter estimate (MP). Standard errors (SE1 to SE5) were generated in WESVAR for each estimated percentage (P1 to P5).
2. The beween- and within-imputation variance for each mean estimated percentage (MP) was computed. The between-imputation variance for each percentage was estimated using the following formula:

$$[(MP-P1)^2 + (MP-P2)^2 + (MP-P3)^2 + (MP-P4)^2 + (MP-P5)^2]/4$$

The within-imputation variance was computed using the following formula:

$$[(SE1^2 + SE2^2 + SE3^2 + SE4^2 + SE5^2)]/5$$

The total imputation variance was computed by summing the between- and within-imputation variances. In doing so, a weight of 1.2 (1 + 1/M, where M is the number of plausible values) was applied to the between-imputation variance. The square root of the total variance provided an estimate of the standard error for the percentage.

B.4 Procedure for Testing Differences between Mean Achievement Scores and Proportions

The approach used to test the significance of differences between mean achievement scores associated with different levels of an explanatory variable involved the following steps:

1. Using the Bonferroni procedure (Dunn, 1961), two-tailed alphas associated with the desired 95% and 90% confidence intervals (i.e., .05 and .10) were divided by the

number of comparisons to be made, and the critical values of t associated with these adjusted alphas were identified *in a statistical table of such values*, using 80 degrees of freedom (the number of variance strata associated with balanced repeated replicate (BRR) method of variance estimation). When two comparisons are made, the adjusted alphas are .025 (.05/2) and .05 (.10/2). For three comparisons, the adjusted alphas are .017 (.05/3) and .033 (.10/3), and so forth.

2. After identifying an appropriate reference category (e.g., female; medium socio-economic status), the differences between each mean and mean of the reference group, and the corresponding standard errors of the difference were computed. The standard error of the difference was computed in Wesvar, so that covariance between groups could be taken into account.
3. 95% and 90% confidence intervals were constructed by adding to and subtracting from each mean difference the product of the corresponding standard error of the difference and the relevant adjusted critical value.

It can be concluded that a difference between a pair of means is not significant if zero falls in the confidence interval around the mean difference. In some cases, a difference which is not significant at the .05 level may be significant at the .10 level. A note on the interpretation of significant differences may be found in Inset 4.3 in Chapter 4.

B.5 Procedure for Computing Correlation Coefficients

Pearson correlation coefficients (*r*) were obtained using the square roots of the coefficients of determination (R^2) associated with each of the five linear regressions computed between the explanatory variable and the response variable (which has five plausible values). Since the distribution of resulting *r*s are not asymptotically normally distributed (they are bounded by +/- 1), each was transformed to a *z*-score using Fisher's transformation (see Schafer, 1997), and the average of the five *z*-scores was then back-transformed to yield a coefficient of correlation. The following formulas were used for transformation and backtransformation:

Transformation: *Fisher's z = ½ log((1+R)/(1–R))*

Back-transformation: *r = (Exp(2*meanz) – 1)/(EXP(2*meanz) + 1)*

where meanz is the mean of the five z-scores.

B.6 Procedure for Calculating Critical Values for Correlation Coefficients

According to Agresti and Finlay (1997), the hypothesis that the population $r = 0$ can be tested using the sample value of *r*, and is equivalent to the *t* test for the hypothesis that $\beta = 0$, where β is the slope of the least square line. While Agresti and Finlay present this in the context of ordinary-least-squares regression with independent observations, it was extended in the current study to complex samples. Since the regression coefficients associated with complex samples and their standard errors can be calculated using software such as Wesvar, the significance of *r* was inferred by computing the *t* statistic (i.e., by dividing the mean β by its standard error), as this also provides a test of linear association in the population. The corresponding *p* value was obtained from a table of critical values of *t*, using 80 degrees of freedom (the number of strata in the BRR method of variance estimation).

Table A4.1. Student Performance on Combined Mathematics, Reading, and Science, by Gender (All Countries)

	Combined Mathematics						*Reading*						*Science*					
	Male		*Female*		*Difference*		*Male*		*Female*		*Difference*		*Male*		*Female*		*Difference*	
	Mean	*(SE)*	*Mean*	*(SE)*	*Diff*	*(SE)*	*Mean*	*(SE)*	*Mean*	*(SE)*	*Diff*	*(SE)*	*Mean*	*(SE)*	*Mean*	*(SE)*	*Diff*	*(SE)*
OECD Countries																		
Australia	522	(2.7)	527	(3.0)	5	(3.8)	545	(2.6)	506	(2.8)	**-39**	(3.6)	525	(2.8)	525	(2.9)	0	(3.8)
Austria	502	(4.0)	509	(4.0)	8	(4.4)	514	(4.2)	467	(4.5)	**-47**	(5.2)	492	(4.2)	490	(4.3)	-3	(5.0)
Belgium	525	(3.2)	533	(3.4)	8	(4.8)	526	(3.3)	489	(3.8)	**-37**	(5.1)	509	(3.5)	509	(3.6)	0	(5.0)
Canada	530	(1.9)	541	(2.1)	**11**	(2.1)	546	(1.8)	514	(2.0)	**-32**	(2.0)	516	(2.2)	527	(2.3)	**11**	(2.6)
Czech Rep.	509	(4.4)	524	(4.3)	**15**	(5.1)	504	(4.4)	473	(4.1)	**-31**	(4.9)	520	(4.1)	526	(4.3)	6	(4.9)
Denmark	506	(3.0)	523	(3.4)	**17**	(3.2)	505	(3.0)	479	(3.3)	**-25**	(2.9)	467	(3.2)	484	(3.6)	**17**	(3.2)
Finland	541	(2.1)	548	(2.5)	**7**	(2.7)	565	(2.0)	521	(2.2)	**-44**	(2.7)	551	(2.2)	545	(2.6)	**-6**	(2.8)
France	507	(2.9)	515	(3.6)	**9**	(4.2)	514	(3.2)	476	(3.8)	**-38**	(4.5)	511	(3.5)	511	(4.1)	0	(4.8)
Germany	499	(3.9)	508	(4.0)	**9**	(4.4)	513	(3.9)	471	(4.2)	**-42**	(4.6)	500	(4.2)	506	(4.5)	6	(4.8)
Greece	436	(3.8)	455	(4.8)	**19**	(3.6)	490	(4.0)	453	(5.1)	**-37**	(4.1)	475	(3.9)	487	(4.8)	**12**	(4.2)
Hungary	486	(3.3)	494	(3.3)	**8**	(3.5)	498	(3.0)	467	(3.2)	**-31**	(3.8)	504	(3.3)	503	(3.3)	-1	(3.7)
Iceland	523	(2.2)	508	(2.3)	**-15**	(3.5)	522	(2.2)	464	(2.3)	**-58**	(3.5)	500	(2.4)	490	(2.4)	**-10**	(3.8)
Ireland	495	(3.4)	510	(3.0)	**15**	(4.2)	530	(3.7)	501	(3.3)	**-29**	(4.6)	504	(3.9)	506	(3.1)	2	(4.5)
Italy	457	(3.8)	475	(4.6)	**18**	(5.9)	495	(3.4)	455	(5.1)	**-39**	(6.0)	484	(3.6)	490	(5.2)	6	(6.3)
Japan	530	(4.0)	539	(5.8)	8	(5.9)	509	(4.1)	487	(5.5)	**-22**	(5.4)	546	(4.1)	550	(6.0)	4	(6.0)
Korea	528	(5.3)	552	(4.4)	**23**	(6.8)	547	(4.3)	525	(3.7)	**-21**	(5.6)	527	(5.5)	546	(4.7)	**18**	(7.0)
Luxembourg	485	(1.5)	502	(1.9)	**17**	(2.8)	496	(1.8)	463	(2.6)	**-33**	(3.4)	477	(1.9)	489	(2.5)	**13**	(3.3)
Mexico	380	(4.1)	391	(4.3)	**11**	(3.9)	410	(4.6)	389	(4.6)	**-21**	(4.4)	400	(4.2)	410	(3.9)	**9**	(4.1)
Netherlands	535	(3.5)	540	(4.1)	5	(4.3)	524	(3.2)	503	(3.7)	**-21**	(3.9)	522	(3.6)	527	(4.2)	5	(4.7)
New Zealand	516	(3.2)	531	(2.8)	**14**	(3.9)	535	(3.3)	508	(3.1)	**-28**	(4.4)	513	(3.4)	529	(3.0)	**16**	(4.2)
Norway	492	(2.9)	498	(2.8)	6	(3.2)	525	(3.4)	475	(3.4)	**-49**	(3.7)	483	(3.3)	485	(3.5)	2	(3.6)
Poland	487	(2.9)	493	(3.0)	6	(3.1)	516	(3.2)	477	(3.6)	**-40**	(3.7)	494	(3.4)	501	(3.2)	**7**	(3.3)
Portugal	460	(3.4)	472	(4.2)	**12**	(3.3)	495	(3.7)	459	(4.3)	**-36**	(3.3)	465	(3.6)	471	(4.0)	**6**	(3.2)
Slovak Rep.	489	(3.6)	507	(3.9)	**19**	(3.7)	486	(3.3)	453	(3.8)	**-33**	(3.5)	487	(3.9)	502	(4.3)	**15**	(3.7)
Spain	481	(2.2)	490	(3.4)	**9**	(3.0)	500	(2.5)	461	(3.8)	**-39**	(3.9)	485	(2.6)	489	(3.9)	4	(3.9)
Sweden	506	(3.1)	512	(3.0)	**7**	(3.3)	533	(2.9)	496	(2.8)	**-37**	(3.2)	504	(3.5)	509	(3.1)	5	(3.6)
Switzerland	518	(3.6)	535	(4.7)	**17**	(4.9)	517	(3.1)	482	(4.4)	**-35**	(4.7)	508	(3.9)	518	(5.0)	**10**	(5.0)
Turkey	415	(6.7)	430	(7.9)	**15**	(6.2)	459	(6.1)	426	(6.8)	**-33**	(5.8)	434	(6.4)	434	(6.7)	0	(5.8)
United States	480	(3.2)	486	(3.3)	**6**	(2.9)	511	(3.5)	479	(3.7)	**-32**	(3.3)	489	(3.5)	494	(3.5)	5	(3.3)
OECD total	*484*	*(1.3)*	*494*	*(1.3)*	*10*	*(1.4)*	*503*	*(1.3)*	*472*	*(1.4)*	*-31*	*(1.4)*	*493*	*(1.3)*	*499*	*(1.3)*	*6*	*(1.5)*
OECD avg	*494*	*(0.8)*	*506*	*(0.8)*	*11*	*(0.8)*	*511*	*(0.7)*	*477*	*(0.7)*	*-34*	*(0.8)*	*497*	*(0.8)*	*503*	*(0.7)*	*6*	*(0.9)*
Partner Countries																		
Brazil	365	(6.1)	348	(4.4)	**16**	(4.1)	384	(5.8)	419	(4.1)	**-35**	(3.9)	393	(5.3)	387	(4.3)	6	(3.9)
HK-China	548	(4.6)	552	(6.5)	4	(6.6)	525	(3.5)	494	(5.3)	**-32**	(5.5)	541	(4.2)	538	(6.1)	-3	(6.0)
Indonesia	358	(4.6)	362	(3.9)	3	(3.4)	394	(3.9)	369	(3.4)	**-24**	(2.8)	394	(3.8)	396	(3.1)	1	(2.7)
Latvia	482	(3.6)	485	(4.8)	3	(4.0)	509	(3.7)	470	(4.5)	**-39**	(4.2)	491	(3.9)	487	(5.1)	-4	(4.7)
Liechtenstein	521	(6.3)	550	(7.2)	**29**	(10.9)	534	(6.5)	517	(7.2)	-17	(11.9)	512	(7.3)	538	(7.7)	**26**	(12.5)
Macao-China	517	(3.3)	538	(4.8)	**21**	(5.8)	504	(2.8)	491	(3.6)	**-13**	(4.8)	521	(4.0)	529	(5.0)	8	(6.8)
Russian Fed	463	(4.2)	473	(5.3)	**10**	(4.4)	456	(3.7)	428	(4.7)	**-29**	(3.9)	485	(4.0)	494	(5.3)	**9**	(4.3)
Serbia	436	(4.5)	437	(4.2)	1	(4.4)	433	(3.9)	390	(3.7)	**-43**	(3.9)	439	(4.2)	434	(3.7)	-5	(3.8)
Thailand	419	(3.4)	415	(4.0)	-4	(4.2)	439	(3.0)	396	(3.7)	**-43**	(4.1)	433	(3.1)	425	(3.7)	-8	(4.2)
Tunisia	353	(2.9)	365	(2.7)	**12**	(2.5)	387	(3.3)	362	(3.3)	**-25**	(3.6)	390	(3.0)	380	(2.7)	**-10**	(2.6)
Uruguay	416	(3.8)	428	(4.0)	**12**	(4.2)	453	(3.7)	414	(4.5)	**-39**	(4.7)	436	(3.6)	441	(3.7)	4	(4.4)

Note: Significant differences are marked in bold.

Table A4.2. Student Performance on the Space and Shape, Change and Relationships, Quantity, and Uncertainty subscales, by Gender (All Countries)

	Space and Shape						*Change and Relationships*						*Quantity*						*Uncertainty*					
	Male		*Female*		*Difference*		*Male*		*Female*		*Difference*		*Male*		*Female*		*Difference*		*Male*		*Female*		*Difference*	
	Mean	*(SE)*	*Mean*	*(SE)*	*Diff*	*(SE)*	*Mean*	*(SE)*	*Mean*	*(SE)*	*Diff*	*(SE)*	*Mean*	*(SE)*	*Mean*	*(SE)*	*Diff*	*(SE)*	*Mean*	*(SE)*	*Mean*	*(SE)*	*Diff*	*(SE)*
OECD Countries																								
Australia	526	(3.2)	515	(2.9)	**12**	(3.9)	527	(3.2)	523	(2.8)	4	(3.8)	518	(2.9)	516	(2.7)	1	(3.7)	535	(3.0)	527	(2.7)	**7**	(3.7)
Austria	525	(4.4)	506	(4.3)	**19**	(5.2)	502	(4.4)	497	(4.4)	5	(5.0)	515	(3.7)	512	(3.7)	3	(4.2)	498	(3.8)	490	(4.0)	8	(4.6)
Belgium	538	(3.2)	520	(3.3)	**18**	(4.6)	539	(3.6)	531	(3.5)	8	(5.1)	530	(3.3)	529	(3.3)	1	(4.7)	529	(3.2)	522	(3.2)	7	(4.7)
Canada	530	(2.1)	511	(2.2)	**20**	(2.5)	546	(2.2)	532	(2.0)	**13**	(2.3)	533	(2.2)	528	(1.9)	**5**	(2.2)	551	(2.2)	538	(1.9)	**13**	(2.3)
Czech Rep.	542	(4.8)	512	(5.1)	**30**	(5.7)	521	(4.5)	508	(4.0)	**13**	(4.9)	531	(4.2)	525	(4.5)	6	(5.1)	509	(3.9)	492	(3.8)	**17**	(4.6)
Denmark	521	(3.4)	504	(3.3)	**16**	(3.7)	520	(3.7)	499	(3.3)	**21**	(3.5)	520	(3.2)	511	(2.9)	**9**	(3.1)	527	(3.4)	505	(3.0)	**22**	(3.2)
Finland	540	(2.6)	538	(2.4)	2	(3.0)	549	(2.8)	537	(2.4)	**11**	(2.8)	550	(2.3)	547	(2.1)	3	(2.3)	551	(2.6)	539	(2.3)	**12**	(2.6)
France	517	(4.3)	499	(3.2)	**18**	(4.7)	522	(4.0)	518	(3.2)	4	(5.0)	508	(3.8)	506	(2.9)	2	(4.4)	512	(3.5)	501	(2.8)	**11**	(4.2)
Germany	506	(4.0)	494	(4.0)	**11**	(4.7)	514	(4.3)	502	(4.4)	**12**	(4.4)	515	(4.2)	514	(3.8)	1	(4.4)	502	(3.9)	484	(3.8)	**18**	(4.0)
Greece	447	(4.7)	428	(3.8)	**19**	(4.0)	445	(5.2)	427	(4.4)	**18**	(4.2)	458	(4.9)	435	(4.0)	**23**	(4.0)	469	(4.3)	449	(3.7)	**20**	(3.7)
Hungary	486	(3.8)	471	(3.9)	**15**	(4.0)	499	(3.6)	490	(3.6)	**10**	(3.9)	497	(3.3)	495	(3.2)	2	(3.6)	493	(3.2)	485	(3.0)	**8**	(3.3)
Iceland	496	(2.4)	511	(2.3)	**-15**	(3.7)	505	(2.4)	514	(2.3)	**-10**	(3.8)	500	(2.5)	528	(2.3)	**-28**	(3.9)	524	(2.4)	532	(2.4)	**-8**	(3.8)
Ireland	489	(3.0)	463	(3.4)	**25**	(4.3)	512	(3.0)	500	(3.5)	**13**	(4.4)	506	(3.1)	497	(3.5)	**9**	(4.3)	525	(3.2)	509	(3.7)	**15**	(4.6)
Italy	480	(4.7)	462	(4.1)	**18**	(6.3)	463	(4.9)	442	(4.0)	**21**	(6.3)	481	(5.0)	469	(4.4)	13	(6.5)	475	(4.5)	451	(3.8)	**24**	(5.9)
Japan	558	(6.3)	549	(4.2)	9	(6.3)	539	(6.4)	533	(4.3)	6	(6.6)	528	(5.6)	525	(3.7)	3	(5.7)	535	(5.6)	521	(3.8)	**14**	(5.7)
Korea	563	(5.1)	536	(6.2)	**27**	(8.0)	558	(4.7)	532	(5.8)	**25**	(7.3)	546	(4.0)	524	(4.9)	**22**	(6.2)	547	(4.1)	525	(5.2)	**22**	(6.6)
Luxembourg	503	(2.2)	474	(2.0)	**28**	(3.3)	494	(2.5)	480	(1.8)	**14**	(3.7)	506	(2.2)	497	(1.6)	**9**	(3.2)	503	(2.2)	481	(1.8)	**22**	(3.5)
Mexico	390	(4.1)	374	(3.5)	**16**	(3.8)	368	(4.9)	360	(4.6)	8	(4.4)	400	(4.8)	388	(4.3)	**12**	(4.5)	392	(3.8)	388	(3.6)	4	(3.5)
Netherlands	530	(3.7)	522	(3.4)	8	(4.3)	554	(3.8)	548	(3.7)	6	(4.3)	526	(4.2)	530	(3.6)	-4	(4.7)	554	(3.6)	544	(3.7)	**9**	(4.1)
New Zealand	534	(2.7)	516	(3.3)	**18**	(3.9)	534	(2.8)	517	(3.4)	**17**	(4.1)	517	(2.7)	505	(3.2)	**12**	(3.9)	538	(2.7)	526	(3.3)	**12**	(3.9)
Norway	486	(3.1)	479	(3.5)	7	(4.3)	490	(3.2)	486	(3.1)	4	(3.3)	494	(2.8)	494	(2.7)	0	(3.3)	518	(3.0)	508	(3.2)	**10**	(3.3)
Poland	497	(3.2)	484	(3.3)	**13**	(3.7)	488	(3.1)	481	(3.4)	**8**	(3.6)	493	(2.9)	491	(3.0)	2	(3.3)	495	(2.8)	492	(2.8)	3	(3.2)
Portugal	458	(4.2)	443	(3.5)	**15**	(3.5)	475	(4.8)	462	(4.0)	**13**	(3.8)	473	(4.1)	459	(3.7)	**14**	(3.3)	476	(4.1)	466	(3.5)	**10**	(3.1)
Slovak Rep.	522	(4.7)	487	(4.1)	**35**	(4.5)	502	(4.1)	486	(3.9)	**16**	(4.2)	519	(4.0)	506	(3.6)	**13**	(3.6)	484	(3.8)	467	(3.4)	**17**	(3.5)
Spain	486	(3.5)	467	(2.4)	**18**	(3.0)	485	(3.8)	477	(2.6)	**8**	(3.3)	495	(3.6)	490	(2.2)	5	(3.1)	493	(3.3)	485	(2.2)	**8**	(2.8)
Sweden	503	(3.0)	493	(3.2)	**10**	(3.5)	506	(3.4)	504	(3.9)	1	(4.3)	515	(2.9)	512	(3.2)	3	(3.6)	515	(3.2)	506	(3.4)	**9**	(3.7)
Switzerland	552	(5.3)	526	(3.7)	**25**	(5.6)	530	(5.1)	515	(3.9)	**15**	(5.3)	536	(4.4)	529	(3.2)	7	(4.6)	526	(4.7)	506	(3.7)	**20**	(5.2)
Turkey	423	(7.6)	411	(6.2)	**12**	(6.0)	425	(9.1)	419	(7.4)	6	(7.2)	421	(8.0)	404	(6.6)	**18**	(6.3)	451	(7.3)	432	(6.1)	**19**	(5.7)
United States	480	(3.3)	464	(3.1)	**15**	(3.2)	488	(3.4)	483	(3.3)	6	(2.9)	478	(3.6)	474	(3.6)	4	(3.4)	493	(3.4)	490	(3.1)	3	(2.8)
OECD total	***494***	***(1.4)***	***478***	***(1.3)***	***16***	***(1.6)***	***493***	***(1.4)***	***484***	***(1.4)***	***10***	***(1.5)***	***490***	***(1.4)***	***484***	***(1.3)***	***6***	***(1.5)***	***497***	***(1.3)***	***487***	***(1.2)***	***11***	***(1.3)***
OECD avg	***505***	***(0.8)***	***488***	***(0.8)***	***17***	***(0.9)***	***504***	***(0.8)***	***493***	***(0.8)***	***11***	***(0.9)***	***504***	***(0.8)***	***498***	***(0.8)***	***6***	***(0.8)***	***508***	***(0.7)***	***496***	***(0.8)***	***13***	***(0.8)***
Partner Countries																								
Brazil	358	(5.2)	343	(4.0)	**15**	(4.1)	344	(7.3)	324	(5.5)	***20***	(4.7)	370	(6.3)	351	(4.8)	**18**	(4.5)	385	(4.9)	369	(3.7)	**15**	(3.4)
HK-China	560	(6.8)	556	(5.0)	4	(6.8)	540	(6.8)	539	(4.8)	1	(7.2)	544	(6.0)	546	(4.1)	-3	(6.1)	564	(6.6)	552	(4.6)	12	(6.7)
Indonesia	369	(3.7)	353	(4.2)	**16**	(2.9)	336	(4.4)	332	(5.4)	4	(3.4)	359	(4.0)	356	(5.0)	2	(3.1)	382	(2.8)	387	(3.4)	**-5**	(2.4)
Latvia	494	(5.2)	480	(3.9)	**14**	(4.2)	487	(5.3)	488	(4.3)	-1	(4.0)	483	(4.4)	480	(3.6)	3	(3.4)	474	(4.2)	474	(3.1)	0	(3.3)
Liechtenstein	557	(7.9)	518	(7.1)	**39**	(12.1)	552	(7.4)	526	(6.5)	**26**	(12.1)	544	(7.0)	523	(5.6)	**21**	(9.9)	538	(6.9)	508	(5.6)	**31**	(10.5)
Macao-China	540	(5.1)	517	(4.3)	**23**	(6.8)	529	(5.0)	509	(4.6)	**20**	(6.6)	542	(4.3)	525	(4.2)	**17**	(6.0)	541	(4.5)	523	(4.2)	**18**	(5.9)
Russian Fed	485	(5.8)	464	(5.0)	**21**	(5.0)	479	(6.0)	475	(4.5)	3	(5.1)	476	(5.0)	469	(4.2)	6	(4.4)	441	(5.1)	432	(3.9)	**8**	(4.2)
Serbia	434	(4.3)	431	(4.9)	3	(4.9)	420	(4.5)	418	(4.9)	1	(4.9)	455	(4.2)	458	(4.7)	-3	(4.7)	431	(4.0)	425	(4.2)	5	(4.2)
Thailand	426	(4.3)	422	(3.8)	5	(4.7)	400	(4.5)	409	(4.0)	-10	(5.1)	412	(4.1)	417	(3.8)	-5	(4.9)	420	(3.4)	425	(3.0)	-5	(4.0)
Tunisia	367	(2.8)	351	(3.2)	**16**	(3.0)	342	(3.0)	331	(3.3)	**11**	(3.0)	372	(2.9)	357	(3.3)	**16**	(2.7)	367	(2.5)	360	(2.8)	**7**	(2.6)
Uruguay	423	(3.6)	402	(3.4)	**21**	(3.6)	420	(4.2)	414	(4.2)	5	(4.4)	436	(3.9)	424	(3.8)	**12**	(4.1)	423	(3.9)	415	(3.6)	**8**	(4.1)

Note: Significant differences are marked in bold.

Table B.3. *Percentages of Irish Students at Each Proficiency Level on the Space and Shape Subscale, and Percentage Differences, by Gender*

Level	Males		Females		All Available	
	Percent	SE	Percent	SE	Percent	SE
<1	8.6	0.85	13.0	1.19	10.7	0.78
1	14.8	1.29	19.0	1.52	16.9	1.15
2	24.4	1.09	26.4	1.37	25.4	0.87
3	24.0	1.24	22.1	1.43	23.0	1.02
4	17.5	1.13	13.3	1.16	15.4	0.77
5	8.3	1.00	5.2	0.83	6.8	0.64
6	2.5	0.43	1.1	0.26	1.8	0.24
Total	100.0		100.0		100.0	

Percentage Differences (Reference Category: Male)

	Difference	SED	BCI95%	
<1	4.4	1.46	**0.5**	**8.3**
1	4.2	1.99	-1.2	9.6
2	2.0	1.75	-2.7	6.7
3	-1.9	1.89	-7.0	3.2
4	-4.2	1.62	-8.6	0.2
5	-3.1	1.30	-6.6	0.4
6	-1.4	0.50	**-2.8**	**-0.05**

Note. SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Table B.4. *Percentages of Irish Students at Each Proficiency Level on the Change and Relationships Subscale, and Percentage Differences, by Gender*

Level	Males		Females		All Available	
	Percent	SE	Percent	SE	Percent	SE
<1	4.7	0.62	5.4	0.75	5.1	0.51
1	10.3	1.04	12.0	1.18	11.2	0.86
2	21.2	1.19	23.9	1.21	22.6	0.84
3	26.8	1.52	27.3	1.23	27.0	1.07
4	22.5	1.13	20.7	1.28	21.6	0.85
5	11.5	0.81	8.9	1.04	10.2	0.63
6	3.0	0.56	1.6	0.47	2.3	0.35
Total	100.0		100.0		100.0	

Percentage Differences (Reference Category: Male)

	Difference	SED	BCI95%	
<1	0.7	0.97	-1.9	3.3
1	1.7	1.57	-2.5	5.9
2	2.7	1.70	-1.9	7.3
3	0.5	1.96	-4.8	5.8
4	-1.8	1.71	-6.4	2.8
5	-2.6	1.32	-6.1	0.9
6	-1.4	0.73	-3.4	0.6

Note. SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted confidence intervals.

Table B.5. Percentages of Irish Students at Each Proficiency Level on the Quantity Subscale, and Percentage Differences, by Gender

Level	*Males*		*Females*		*All Available*	
	Percent	*SE*	*Percent*	*SE*	*Percent*	*SE*
<1	5.1	0.65	6.0	0.77	5.6	0.57
1	11.4	1.16	13.2	0.98	12.3	0.85
2	22.7	1.25	23.4	1.31	23.0	1.00
3	26.6	1.77	27.1	1.12	26.9	1.06
4	21.2	1.28	20.0	1.42	20.6	0.84
5	10.3	0.83	8.6	1.04	9.5	0.62
6	2.7	0.47	1.8	0.50	2.2	0.36
Total	100.0		100.0		100.0	

Percentage Differences (Reference Category: Male)

	Difference	*SED*	*BCI95%*	
<1	0.9	1.01	-1.8	3.6
1	1.8	1.52	-2.3	5.9
2	0.7	1.81	-4.2	5.6
3	0.5	2.09	-5.1	6.1
4	-1.2	1.91	-6.3	3.9
5	-1.7	1.33	-5.3	1.9
6	-0.9	0.69	-2.7	0.9

Note. SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted confidence intervals.

Table B.6. Percentages of Irish Students at Each Proficiency Level on the Uncertainty Subscale, and Percentage Differences, by Gender

Level	*Males*		*Females*		*All Available*	
	Percent	*SE*	*Percent*	*SE*	*Percent*	*SE*
<1	3.3	0.52	4.0	0.65	3.6	0.45
1	9.2	0.99	11.3	1.14	10.2	0.74
2	19.3	1.25	23.0	1.29	21.2	0.86
3	26.5	1.34	26.5	1.18	26.5	0.93
4	22.7	1.27	21.3	1.36	22.0	0.93
5	13.6	0.96	11.2	1.21	12.4	0.72
6	5.2	0.65	2.7	0.51	4.0	0.39
Total	100.0		100.0		100.0	

Percentage Differences (Reference Category: Male)

	Difference	*SED*	*BCI95%*	
<1	0.7	0.83	-1.5	2.9
1	2.1	1.51	-2.0	6.2
2	3.7	1.80	-1.1	8.5
3	0.0	1.79	-4.8	4.8
4	-1.4	1.86	-6.4	3.6
5	-2.4	1.54	-6.6	1.8
6	-2.5	0.83	**-4.7**	**-0.3**

Note. SED = standard error of difference; CI95L, CI95U = Bonferroni-adjusted confidence intervals. Confidence intervals for significant differences (p ≤ .05) are highlighted in bold.

Appendix C

C.1 Brief Description of the Main Changes to the Mathematics Curriculum at Primary Level (Implemented in 2002)

The revised primary school mathematics curriculum (Department of Education and Science, 1999) represents quite a radical change from its 1971 predecessor. It focuses more on problem-solving and real-life contexts. The Junior Certificate Mathematics Teacher Guidelines (2002) state that

> …students emerging from the revised [primary mathematics] curriculum should be more likely than their predecessors to look for meaning in their mathematics and less likely to see the subject almost totally in terms of the rapid performance of techniques. They may be more used to active learning, in which they have to construct meaning and understanding for themselves. (2002, p. 5)

The changes in emphasis and content, which are consistent with the PISA approach, are summarised in Table C.1.

Table C.1. Summary of Changes to Primary School Mathematics Curriculum (1971 and 1999)

Changes in emphasis	*Changes in Content*
More emphasis:	*New areas:*
real-life contexts	Calculators introduced at fourth class level
hands-on activities	Increased emphasis on estimating
relational and instrumental understanding	Increased coverage of data handling
appropriate use of mathematical language	Introduction to probability
recording	Subtraction using the decomposition method
problem-solving skills	
	Areas excluded:
Less emphasis:	Unrestricted calculations
routine procedures with no context	Subtraction of negative integers
complicated calculations	Formal treatment of lowest common multiple and highest common factor
	Use of formulae that students have not developed
	Two-step equations and rules
	Sets (except in developing concept of number)
	π and advanced properties of circles

Source. Junior Certificate Mathematics Teacher Guidelines (2002, pp. 18-19).

C.2 Detailed Description PISA 2003 Test-Curriculum Rating Project Rating Scales

Concept: The concept scale requires raters to read through the text of the question and rate how familiar they would expect the typical third year student to be with the *concept underlying the question*. Note that by 'concept' here, we mean *a mathematical principle in its abstract form.*[61] Thus raters are asked to identify the *abstract mathematical concept* underlying the item and *not* to concern themselves with its application for the concept scale. Raters are also asked to identify the *specific* mathematical concept underlying the item rather than at a more general level because this will make it easier to locate it within the mathematics curriculum implemented in Ireland. In the event that multiple and interlinked

[61] This is in contrast to *the demonstration of understanding*, i.e. *the application of a mathematical principle in a specific instance*.

underlying concepts are identified, which can often happen in the case of mathematics, a pragmatic and holistic approach is to be taken. That is, if one concept, more so than others underlying a particular item, is deemed essential to respond to or find the answer to an item, then precedence should be given to that concept in particular in rating the concept familiarity of the item. The question to be answered (for each syllabus level) is:

How familiar would you expect the typical third year student to be with the specific mathematical concept(s) underlying this item?

Context/Application: The context/application scale requires raters to consider the stimulus text and the question and to rate how familiar they would expect the typical third year student to be with *applying the concept(s) underlying the question* in the type of *context* suggested by the question and stimulus text. Context can be conceptualised in a number of different ways, but the focus is at a fairly general level, i.e. whether students are familiar with the mathematical concept(s) being contexualised in this way, and whether the contextualisation of the question would be likely, based on the syllabus or Junior Certificate Examination, to guide them to (or distract them from) the successful application of the concept. The question to be answered (for each syllabus level) is:

How familiar would you expect the typical third year student to be with the application of the specific mathematical concept(s) underlying this item in the type of context suggested by the item and stimulus text?

Format: The format scale requires raters to read through the text of the question and to rate how familiar they would expect the typical third year student to be with *applying the concept(s) in the type of format in which the question and accompanying stimulus text is presented.* Format refers to the *layout* of an item, the *question type* (e.g., multiple-choice, free response), and as such is distinct from context. For example, some PISA mathematics items may contain a long text passage; however, Irish students, who are used to more ‘stark’ exposition of mathematics problems, would not be familiar with a rich text-based format or presentation. The question to be answered (for each syllabus level) is:

How familiar would you expect the typical third year student to be with the application of the specific mathematical concept(s) underlying this item in the type of format suggested by the item and stimulus text?

Table C.2. Percentage of New PISA 2003 Mathematics Items on Which There Was a Lack of Consensus (N=65)

	Higher	*Ordinary*	*Foundation*
Concept	33.8	29.2	33.8
Context	46.2	15.4	13.8
Format	47.7	20.0	20.0

C.3 Description of a Sample of Items on the 2003 Junior Certificate Mathematics Examination with Reference to PISA 2003 Mathematics

Figure C.1 shows a small sample of questions taken from the 2003 Junior Certificate mathematics examination papers. The items were chosen deliberately to contrast with the sample items from the PISA 2003 mathematics assessment shown in Appendix A. The selection does not include many of the Junior Certificate areas, e.g., trigonometry, graphs and functions, transformation geometry; and the items are drawn mainly from the applied arithmetic and measure, basic algebra, and statistics strands[62].

The first question in Figure C.1 (Higher level, algebra) involves the solution of three equations (one linear, one quadratic, and one involving the simplification of an expression with a quadratic term) with no context, and students are presumably drawing procedures from memory and applying them in a routine manner. One might contrast this question to the PISA unit Walking (M124; Appendix A), where students are presented with a simple algebraic equation that describes the relationship between pacelength and number of steps per minute. They must apply this equation in the first question to find pacelength, and in the second, they must compute walking speed in both metres per minute and kilometres per hour. This is a non-routine procedure that draws, in the second question, on skills from applied arithmetic and measure as well as those associated with algebra. Irish students did comparatively poorly on both items: for the first item, 34.7% of Irish students compared to 21.8% of students across the OECD succeeded in getting the first item partially correct; 22.9% of Irish students compared with 36.3% succeeded in getting the item fully correct. This suggests that Irish students were able to substitute the correct numbers in the formula but failed more often than OECD students in general to manipulate the formula to find the solution. The third item which required conversion to both metres per minute and kilometres per hour, is also a partial credit item, and the response pattern suggests that Irish students were about as likely as students across the OECD (around 20% in each case) to substitute figures into the formula but again failed to manipulate the formula or convert to kilometres per hour. Just 8.5% of Irish students compared to 17.0% of students across the OECD were able to progress further with the problem to give a more complete answer.

The second question in Figure C.1 (Higher level, measure) again presents information without a real-life context and without any redundant information. Although, computationally, this question is moderately challenging, one can see that the computations required are of a routine nature. The PISA unit Staircase (M547, Appendix A) is similar to this Junior Certificate question in that it presents no redundant information (apart from the measure of the base, which is irrelevant) and directs students clearly to find the height of each step. Across the OECD, 78.0% of students correctly figured out the answer; in Ireland, 79.7% of students did so. The PISA unit Carpenter (M266) is more challenging in that it requires students to compute the perimeter of four shapes and compare their answers to a given amount. Across the OECD, just 20.0% of students correctly worked out the perimeter of all four shapes; 13.0% of Irish students did so.

The third question in Figure C.1 (Higher level, statistics), asks students in the first part to compute the mean of a series of numbers; and, given one unknown in a series of numbers plus the mean, to find the unknown. These questions can be compared to the PISA unit Test Scores (M513, shown in Appendix A) which shows a grouped frequency distribution for two groups on one graph. Students must construct a mathematical counterargument to the statement that one group is doing better than the other. To do so, students can choose from three arguments: that more students in the lower group pass the test, that if you remove the one outlier from the lower group their mean score is higher, or that the group with the overall lower score contains more higher-scoring students. Irish students demonstrated a comparatively good understanding of this question, with 40.8% of them getting the item correct, compared with 32.2% of students across the OECD as a whole.

---

[62] The full set of examination papers can be obtained from http://www.examinations.ie

The fourth question in Figure C.1 (Ordinary level, applied arithmetic), requires students to convert an amount of yen to euro given the exchange rate. This is similar to the PISA unit Exchange Rate (M413, shown in Appendix A). The first problem in the unit requires students to convert an amount of Singapore dollars to South African Rand, given the exchange rate. Irish students did well on this straightforward calculation (83.2% correct compared with 79.7% across the OECD). The second item in the unit requires students to apply the concept of changing rates: given a new exchange rate, they are asked to convert an amount of South African currency into Singapore dollars. Irish students again performed quite well on this item (76.3% correct in Ireland compared to 73.9% correct). The third question requires students to compare the first and second exchange rates. This was a more difficult item, and 40.8% of Irish students compared to 40.3% across the OECD got the answer correct.

The first two parts of the fifth question in Figure C.1 (Ordinary level, measure) involve the application of the speed/distance/time equation; the third part is more challenging, involving the combination of several pieces of information. The level of sophistication required here is comparable to the computation involved in the Walking unit described above, except that the Walking unit requires manipulation of an algebraic equation as well as computation of speed and conversion of units.

The sixth question in Figure C.1 (part B, ordinary level, statistics), involves the manipulation of pie chart data. Students need to be adept at translating between degrees, number, and percent. This can be compared to the PISA unit Exports (M438; Appendix A), which also involves the interpretation of a pie chart, but in the second question, in combination with a frequency distribution; i.e., students must combine the information in the pie chart and the graph to indicate the value of one export. The question requires conversion of a percentage to an amount of money. In Ireland, 50.8% of students managed to do so, compared to 48.3% across the OECD.

The next question in Figure C.1 (Foundation level, applied arithmetic), requires students to compute the total cost of four items. This question is similar to the PISA unit Skateboard (M520, Appendix A) where students are shown a list of costs for skateboard parts. The first question in this unit asks students to compute the minimum and maximum prices for a self-assembled skateboard. Irish students did well on this item, with 69.0% correctly computing both the lowest and highest prices (across the OECD 66.7% of students did so). The second item requires students to indicate the number of different skateboard designs. Irish students did very poorly on this item, with only 30.2% giving the correct response; 45.5% of students across the OECD got the item correct. The third item requires students to indicate, given a certain budget, how much can be spent on each skateboard part. Across the OECD, 49.8% of students got this item correct, and it was 50.3% in the case of Ireland.

The last question in Figure C.1 (Foundation level, algebra) requires students to solve two equations in the format $Nx + y = z$ and give values for y for $x = 1, 2, 3$ and 4, given $y = 2x - 1$. Hence this is quite basic algebra. This question might be compared to the PISA unit Walking, discussed above: the equation showing the relationship between speed and pacelength is also quite basic but the manner in which students are expected to manipulate it and apply it is more complex.

Figure C.1. Selection of Sample Questions from the 2003 Junior Certificate Mathematics Examination

Higher Level
Paper 1 Question 3

A Given that $p = \frac{x + 2y}{p}$, express x in terms of y and p.

B
(i) Multiply out: $(3x - 1)(2x^2 + x - 4)$.
(ii) Evaluate your answer to part (i) when $x = -2$.

C
(i) Solve $x^2 = 13x + 36 = 0$.
(ii) Hence, find the two values of $t \in \mathbf{R}$ for which

$$\left(\frac{1}{t} + 2\right)^2 - 13\left(\frac{1}{t} + 2\right) + 36 = 0.$$

Paper 2 Question 1 (Part B)
A solid rectangular metal block has length 12 cm and width 5 cm. The volume of the block is 90 cm^3.

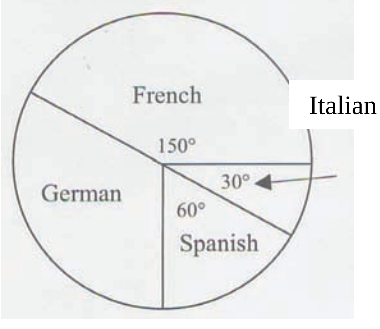
(i) Find the height of the block in cm.
(ii) Find the total surface area of the block in cm^2.
(iii) Each cm^3 of metal has mass 8.4 g. The total mass of a number of these metal blocks is 113.4 kg. How many blocks are there?

Paper 2 Question 6 (Part A)
(i) Show that 13 is the mean of the numbers 6, 11, 15, 16, 17.
(ii) 14 is the mean of the numbers 6, 11, 15, 16, 17, x. Find the value of x.

Ordinary Level
Paper 1 Question 2 (Part B)
(i) 1 euro = 120 Japanese yen. Change 3000 yen to euro.

Paper 2 Question 1 (Part B)
A person travels 48 km to work in the morning and travels home by the same route in the evening.
(i) It takes 45 minutes to travel to work. Calculate the average speed in km/hr.
(ii) The person returns home at an average speed of 72 km/hr. How many minutes does the journey home take?
(iii) At what time should the person leave work at in order to arrive home at 20:15?

Paper 2 Question 3
B
Each student in a class studies one of four languages: French, German, Spanish and Italian. The pie-chart represents the number of students that study each language.



(i) What is the measure of the angle for German?
(ii) 10 students study French. How many students study Italian?
(iii) How many students are in the class?
(iv) How many students do not study Spanish?

Figure C.1. Selection of Sample Questions from the 2003 Junior Certificate Mathematics Examination (continued)

Ordinary Level
Paper 2 Question 3
C
The following gives the number of days that each of 30 pupils was absent during May:

1	0	2	3	1	0	0	4	5	5
6	5	3	2	0	5	1	0	4	5
3	2	3	6	5	4	3	6	6	0

(i) Complete the following frequency table:

Number of days absent:	0	1	2	3	4	5	6
Number of pupils:							

(ii) Compute the mean number of days absent per pupil during May.
(iii) What percentage of the pupils were absent for three days or more?

Foundation Level
Question 1 (Part C)
Find the total cost of

One bus ticket @ 8.00
One CD @ 13.50
Two concert tickets @ 15.60 each
Two tee shirts @ 8.50 each.

[Answer space leads students to the method of solution, by prompting multiplication and totalling.]

Question 5

A Find the value of $3x + 2$ when $x = 4$.
B
(i) Solve for x: $x + 5 = 12$
(ii) Solve for x: $3(x - 1) = 9$
C
(i) Given that $y = 2x - 1$, complete the table below:

x	1	2	3	4
y				

(ii) Draw the graph of $y = 2x - 1$, from $x = 1$ to $x = 4$. [a graph area is provided as an answer space].

Note. The layout of the questions is more compact than the original appearing in the examination papers. In some instances, only some sub-parts of a particular question are shown for reasons relating to space limitations. For the complete papers, see http://www.examinations.ie.

Index:
Explanations of Statistical Terms and Procedures

ISBN 0-900440-19-8

9 780900 440199

Educational Research Centre,
Dublin 9.
http://www.erc.ie