SCOTTISH EXECUTIVE
Education Department

Programme for International Student Assessment (PISA) 2003

Initial Report on Scotland's Performance in Mathematics, Science and Reading

Programme
for
International

Student Assessment (PISA)
2003

Initial Report on Scotland's Performance in Mathematics, Science and Reading

Graham Thorpe

The SCRE Centre,
University of Glasgow

This report is a publication of the research group within Information, Analysis and Communication Division, which is responsible for providing analytical services within the Scottish Executive Education Department (SEED). Their work is part of a multidisciplinary unit (consisting of researchers, economists and statistics staff) and the staff undertakes and funds economic analysis and social research in the fields of: school education; children, young people and social work: architecture; and tourism, culture and sport.

Acknowledgements
The author would like to thank Liz Levy, Information, Analysis & Communication Division (IAC) for editing, checking data and providing some additional analysis; and Rik Kay (IAC) for assembling and formatting the report. Both the SCRE Centre and the Scottish Executive Education Department express their gratitude to the schools and individuals who took part in the PISA 2003 as mentioned in the Introduction.

A limited number of additional copies may be obtained from the Dissemination Officer, Information, Analysis and Communication Division, Scottish Executive Education Department, Victoria Quay, Edinburgh EH6 6QQ (telephone 0131-244-0316) or e-mail recs.admin@scotland.gsi.gov.uk. Copies of this report and our other education research publications can be downloaded from the main Scottish Executive website "Publications" page and also from our web section for education research www.scotland.gov.uk/insight/

Contents

Executive summary

Introduction

The OECD Programme for International Student Assessment (PISA) assesses the performance in reading literacy, mathematical literacy and scientific literacy of 15 year old students in its member countries.

PISA examines young people's ability to use their knowledge and skills to meet real-life challenges rather than whether they have mastered a specific school curriculum.

In PISA 2003, the main subject assessed was mathematics, with reading and science forming the minor domains.

Scotland took part in PISA 2003 as an independent National Centre, meaning that it participated fully in all PISA activities as though it were a full country including separate quality monitoring and adjudication of test administration.

Over 275,000 students took part in PISA 2003 from 41 participating countries (all 30 OECD countries and 11 non-OECD ‘partner’ countries). Although the United Kingdom as a whole failed to meet the rigorous sampling criteria for the study (meaning that its results are not considered sufficiently reliable to be reported in full) Scotland and Northern Ireland did, separately, meet the criteria and their results are reported in an annex to the international report.

The intended sample for Scotland was of 108 schools and 35 pupils in each of these. The school sample was a stratified, random sample representative of all mainstream secondary schools in Scotland (local authority, grant-aided, and private schools). Of the 108 Scottish schools that were recruited to the study, 98 returned completed tests and questionnaires, a response rate of 90%. Just over 2,700 students completed tests and questionnaires.

This report presents results for Scotland and the 29 OECD countries that fully met the OECD criteria for acceptability. It does not report on results from the ‘partner countries’.

This present report is concerned with the main set of test results from the study. Further themed reports planned for 2005 are likely to report on the questionnaire data gathered from schools and pupils, the relationships between these data and the results obtained in mathematics, and the domain of problem solving (which was additional in PISA 2003).

Student Proficiency in Mathematical Literacy

Scotland’s mean score in mathematics is 524. This is significantly above the OECD average. Eight OECD countries have mean scores higher than Scotland but only Finland, Korea, and The Netherlands have mean scores that are significantly higher.

In mathematics literacy, male students outperformed female students in all countries except Iceland. For the 29 OECD countries as a whole, the average difference is close to 11 points. In Scotland it is 7 points.

PISA 2003 reports on student performance at six ‘proficiency levels’ – descriptions of the kind of mathematical competency demonstrated by students. Just below one-third (31%) of students across the OECD area as a whole performed at the top three proficiency levels (Level 4 or above) and 3.5% at the highest level, Level 6. In Scotland, 41.2% of students attained Level 4 or better and 3.9% achieved Level 6.

As well as assessing overall performance in mathematics, PISA 2003 assessed pupil performance in 4 ‘content areas’: *Space and shape* (which draws upon the discipline of geometry); *Change and relationships* (which relates most closely to algebra); *Quantity* (which covers those aspects of mathematics bearing upon number); and *Uncertainty* (which lies within the area of probability and statistics). The mean scores for Scotland in each of the four content areas are significantly higher than those for the OECD as a whole.

Both in respect of the OECD taken as a whole and relative to the individual countries that constitute the OECD, Scottish students do very well in the content area of *Uncertainty*. Only one OECD country had a mean score in this area that was significantly higher than that of Scotland. *Change and Relationships* is the next strongest area. Only two OECD countries have higher mean scores than Scotland. In the other two content areas, Scotland's performance is good but not strong (nine countries have a mean score that is significantly higher than that of Scotland's in *Space and Shape* and five countries have a mean score that is significantly higher in *Quantity*).

In each content area, male students attain higher mean scores than females both in the OECD as a whole and in Scotland.

It is not possible to compare overall performance in mathematics in PISA 2000 and 2003 (for any participating country). The only valid comparison is confined to the two content areas of *Space and Shape* and *Change and Relationship*. Performance could be compared across 25 OECD countries and Scotland.

The average score in *Space and Shape* across these 25 countries as a whole did not change significantly. On a common scale, the mean score in PISA 2000 was 494, while that in PISA 2003 was 496 points. Scotland showed a slight, though non-significant drop in performance, with a mean score in PISA 2000 of 511 points and in PISA 2003, of 507 points.

The average score in *Change and Relationship* showed a significant increase. On a common scale, the mean score in PISA 2000 was 488, while that in PISA 2003 was 499 points. Although Scotland's mean score in this content area (529) is one of the highest in PISA 2003, this mean decreased since PISA 2000 by 1 point.

PISA 2003 provides encouraging evidence that Scotland is succeeding in 'closing the gap in attainment'. One of the ways in which PISA measures equity in attainment is by comparing the score gap between students at the 75th percentile level and those at the 25th percentile. The smallest gap (115 points) is found in Finland. Scotland has the same score gap as Finland.

Student Proficiency in Reading Literacy

Scotland's mean score of 516 is (statistically) significantly above that for the OECD as a whole. Only the three top performing OECD countries on the PISA 2003 reading literacy scale (Finland, and Korea and Canada) had mean reading scores that were significantly higher than that of Scotland.

In every OECD country the mean reading score for female students is significantly higher than that for male students. On average, across the OECD, the difference in performance between female and male students is 34 points. In Scotland, the difference was 24 points.

Overall, 55% of students in the OECD attained the top 3 proficiency levels in reading literacy (Levels 3,4 and 5) and 8% attained the highest level, Level 5. In Scotland, 68% performed at the top three levels and 9% at Level 5.

Between PISA 2000 and 2003 mean score in reading literacy in Scotland dropped by 11 points - a drop that is of borderline statistical significance (i.e. significant at the 10% level of significance but not at the 5% level).[1]

[1] The level of significance refers to the odds that the difference between the scores is due to there being an actual difference, and not to chance. At the 5% level there is a 1 in 20 chance that the difference is not a true difference (i.e. due to chance), while at the 10% level these odds are greater, a 1 in 10 chance that the difference is not a true difference.

Although there has been a drop in Scotland's mean score, this drop has not been experienced across all ranges of ability. Despite the trend across the OECD as a whole, the performance of our lowest 25% of students has improved (though not significantly).

Scotland's score gap in reading between students at the 75th percentile level and those at the 25th percentile level is 116. Only 4 countries had a score gap that was narrower than this. The smallest gap (105 points) is again found in Finland.

In PISA 2000, Scotland's score gap (between the 25th and 75th percentiles) in reading was 136 points. This means that between 2000 and 2003, our score gap narrowed by 15%. No other OECD country narrowed its score gap by as much as this.

Student Proficiency in Scientific Literacy

Scotland's mean score in science literacy (514) is significantly above the OECD average. Of the eight OECD countries with higher mean scores than Scotland, only the top three countries (Finland, Japan and Korea) can be said with certainty to have students who do better, on average, than those in Scotland.

While in general male students did better than females in the OECD at large, this is not universally the case. In Scotland there is a difference of 8 points in favour of male students, but this difference is not statistically significant.

Scotland's mean science score fell between PISA 2000 and PISA 2003, but the 8 point drop, from 522 to 514 points, is not statistically significant.

The drop in performance in Scotland was more or less uniform across the whole ability range. Performance by students at the very lowest levels of ability, those at the 5th percentile and 10th percentile levels fell by approximately 5 points, while those at the 25th 75th 90th and 95th percentile levels fell by between 7 and 9 points.

In Scotland, the score gap between students at the 75th percentile level and those at the 25th level was 140 points. Scotland's rating on this equity scale is neither particularly good nor particularly bad, compared with the OECD countries.

In PISA 2000, Scotland's score gap for science was 142. This means that between 2000 and 2003 we narrowed the gap by only 2 score points. However, only three OECD countries managed to narrow their gap in science between these two sweeps of PISA: Belgium (by 1 score point); Denmark (by 3 score points) and Hungary (by 8 score points). In all other OECD countries the gap increased.

Chapter 1: Introduction

The OECD established its Programme for International Student Assessment (PISA) to assess the performance in reading literacy, mathematical literacy and scientific literacy of 15 year old students in its member countries. The results from the studies contribute school outcome measures to the OECD's educational indicators programme[2]. Since PISA's inception its scope has broadened to include non-OECD countries as well. These non-OECD countries are referred to as "partner countries".

The first PISA survey took place in 2000[3], the second in 2003, and work has started on the third, to take place in 2006. Each study focuses in turn upon one of the above three literacies, while allocating a subsidiary role to the other two. Consequently, the whole series has a major cycle of nine years. In PISA 2000 the majority of assessment items were in reading literacy and the minority in mathematics and science literacy, while in PISA 2003 the majority are in mathematical literacy, the other two domains playing the minor roles. In PISA 2006 the majority are in scientific literacy. To provide continuity between studies and a measure of change over the three-year intervals, a proportion of the test items from previous studies are re-used in subsequent ones.

The scope of PISA has grown. In 2003 an additional, subsidiary, domain of problem solving was added to the other three, though this is not to be carried forward to 2006. (The results of the problem-solving tests are not reported in this initial report but will be reported on in 2005.) A trial of computer-based assessment will form part of PISA 2006.

The domains covered by PISA are defined in terms of:

- The content or structure of knowledge that students need to acquire
- The processes that need to be performed
- The contexts, or situations, in which knowledge and skills are applied.

PISA examines young people's ability to use their knowledge and skills to meet real-life challenges rather than whether they have mastered a specific school curriculum. For instance, PISA defines reading literacy as the ability to understand, use and reflect on written texts in order to participate effectively in life. Its concern is not one of making a parallel assessment to those provided by school examinations.

PISA provides a broad assessment of comparative learning outcomes towards the end of compulsory schooling to guide policy decisions and provide insights into factors that contribute to the development of knowledge and skills, and the extent to which these factors are common to different countries.

Design of PISA 2003

In 2003 over 275,000 students took part in PISA from 41 participating countries[4]. All 30 member countries of the OECD participated and 11 non-OECD, partner, countries. One OECD country, the United Kingdom[5], failed to meet the rigorous sampling criteria for the study, meaning that its results are not considered sufficiently reliable to be reported in full. Scotland and Northern Ireland did, separately, meet the criteria. Consequently, this report

---

[2] See OECD (OECD, 2003)

[3] See OECD (OECD, 2001)

[4] Note that the OECD counts full national entities as "countries". Thus the UK is counted once only in this figure of 41, although England with Wales, Scotland, and Northern Ireland participated as three independent units within the actual study. Similar situations held in other "countries".

[5] Appendix A gives more information on which sampling criteria the UK failed to meet.

gives results for Scotland and 29 OECD countries only. The UK and the partner countries are not included in this national report.

Pencil and paper assessments were used to assess students, with two hours of assessment for each student. Various questions types were used, including multiple choice questions and questions requiring students to construct their answers. Each student's particular assessment tasks were drawn from a total of six and a half hours of assessment items, with different students taking different combinations of items from this pool. This combination spanned all the topics being assessed in the 2003 survey, namely: mathematics, reading, science and problem solving. Students also completed a 30 minute questionnaire and senior teachers completed a questionnaire about their schools. Test sessions were supervised by external administrators in most countries, including Scotland.

PISA 2003 International Report

The OECD is to publish the first results from the 2003 study in *'Learning for Tomorrow's World – First Results from PISA 2003'*, (OECD, 2004). The results will also appear in the OECD's annual compilation of educational indicators *Education at A Glance.* Subsequently, a number of thematic reports will be produced. These will primarily concentrate on the findings from the mathematics section of the study. All these reports can be found on the OECD web-site at "http://www.oecd.org/".

Following OECD convention, the report gives two OECD values for the various results it tabulates, eg, two values for the overall mean scores are tabulated, not just the single value that might be expected. The first of the two values (the OECD average) is the average of each of the individual country averages. The second value (the OECD total) is weighted to take account of the different numbers of 15 year-old students in each of these countries. Thus, while the USA contributes equally to the first value, it makes a much larger contribution to the second. Iceland on the other hand while also contributing equally to the first value, contributes much less to the second. The first index is useful for comparing how a country compares with a typical OECD country, the second for comparing the performance of students in any one country with all students in the OECD. We have tried to ensure that, in determining which value to use in the following analysis, we are consistent with the approach that is used in the PISA 2003 international report.

PISA in Scotland

Scotland is not a member country of the OECD, but a part of the United Kingdom. For the PISA 2003 study it operated as an independent National Centre, meaning that it participated fully in all PISA activities as though it were a full country. This allowed for greater input into the planning stages of the programme as well as separate quality monitoring and adjudication of test administration. Structural differences in secondary schooling between Scotland and the other parts of the United Kingdom were more readily managed by treating Scotland as a distinct National Centre, and the results were more readily analysed by the International Study Centre as a consequence. Additionally, results for Scotland will appear in tables in the annex to the international report on the study, enabling the standing of Scotland to be readily compared with that of other countries, unlike PISA 2000, where results were listed for the whole of the UK only.

The intended sample for Scotland was of 108 schools and 35 pupils in each of these. Similar numbers were required for every participating country. The school sample was a stratified, random sample representative of all mainstream secondary schools in Scotland (local authority, grant-aided, and private schools). Special schools were excluded. The pupil sample spanned S4 and S5, as 15 year old pupils are found in both stages, and the required pupils were selected randomly by date of birth. All pupils except those with any of a defined set of severe learning difficulties, and those who explicitly refused to participate or whose parents refused for them, were eligible for assessment.

Of the 108 Scottish schools, 98 returned completed tests and questionnaires, a response rate of 90%. Just over 2,700 students completed tests and questionnaires. These response rates met the OECD's strict criteria for acceptability. The SCRE Centre and the Scottish Executive Education Department is very grateful to the schools and pupils that agreed to take part in PISA 2003. We understand the many demands that are placed on schools and on 15 year olds and that participation in PISA creates some disruption at a time when many are preparing for exams. We are pleased that schools see PISA as an important and worthwhile study and hope that this initial report and, in particular, the further reports planned for 2005 helps highlight the value of participation.

The Scottish Report

As already noted, this report presents results for Scotland and the 29 OECD countries that fully met the OECD criteria for acceptability. Comparative results for the UK are not given, as the OECD considered these too unreliable (see Annex A for explanation). No results for the 11 non-OECD partner countries are given: readers wishing to know about any of these countries should refer to the main, international reports. Similarly, those wishing to know how Northern Ireland fared should refer to its national report.

The decision to limit the report to Scotland and the OECD was one of expediency. To have reported results for all 41 countries, would have meant excessively long tables and charts. As the 29 OECD countries included most countries whose results would be of interest to a Scottish readership, including most European Union countries, this was the best compromise. It also seemed to be an approach that was consistent with the presentation of the results in the international report where, as already mentioned, international averages are based on the performance of OECD countries only.

This present report is concerned with the main set of test results from the study. Further themed reports planned for 2005 are likely to report on the questionnaire data gathered from schools and pupils, the relationships between these data and the results obtained in mathematics, and the domain of problem solving (which was additional in PISA 2003).

This first Scottish report has been written at the same time as the international report and has drawn heavily on the results presented in that report. Clearly, one unavoidable consequence of this is that it may not be possible to carry through to the Scottish report all the last minute editorial revisions made to the international report. Some of the values reported in the Scottish report may, consequently, differ from those in the published International report. Any such discrepancies should be slight and should not change any conclusions drawn about the Scottish performance.

Trends in Mathematics and Science Study (TIMSS) 2003

At the same time as taking part in PISA 2003, Scotland also took part in the Trends in Mathematics and Science Study (TIMSS) which is run by The International Association for the Evaluation of Educational Achievement (IEA). TIMSS measures mathematics and science performance at P5 and S2. This study reported on 14th December 2004 and the results show that our performance in mathematics is not significantly different from the international average at P5 but is significantly higher than this average at S2. In science, our performance was significantly above the international average at both P5 and S2.

When drawing analysis between the TIMSS and PISA results one must take into account the differences in the nature of the PISA and TIMSS approach. The assessments used in TIMSS were constructed on the basis of an analysis of the intended curriculum in each participating country so as to cover the core material common to the curriculum in the majority of participating countries. By comparison the assessment material used in PISA cover the range of skills and competencies that were considered to be crucial to an individual's capacity to fully participate, in respect of the assessment domains, in a successful modern society (in other words, to apply their skills in 'real life' situations). The other important difference between PISA and TIMSS is the stage at which students are assessed. While PISA assesses pupils towards the end of compulsory schooling (and therefore assesses the cumulative impact of schooling), TIMSS assesses pupils in middle primary and in early secondary.

For further detail about TIMSS and Scotland's performance in this study, see http://timss.bc.edu/.

Domestic Survey of Attainment

The results of PISA 2003 are also more encouraging that those of Scotland's domestic Assessment of Achievement Programme (AAP). In 2003, AAP examined student attainment in science. While PISA reveals that (although not our strongest subject), towards the end of compulsory schooling, students in Scotland perform relatively well in science compared with their international peers. AAP suggests that our younger secondary school students are not reaching the attainment targets in science that are set nationally. At secondary school level, AAP tests a sample of pupils from S2. The results show that fewer than 10% of this sample was 'secure' at the target attainment level (Level E) and a high proportion failed to show even 'basic' attainment at this level.

Chapter 2: Student Proficiency in Mathematical Literacy

How Mathematical Literacy is Defined

The following paragraphs briefly outline the definition and assessment of mathematics literacy used for PISA. A full account is to be found in *The PISA 2003 Assessment Framework,* 2003 (OECD) and in Chapter 2 of the international report on the PISA 2003 survey results referred to earlier.

Centrally, the PISA study is predicated upon an understanding of mathematics as now being of significance for all adults in society rather than just for a minority of specialists as has been assumed, either overtly or implicitly, in previous studies. The concern is consequently not with how well students confront problems designed specifically to assess the concepts and techniques taught at school but whether they can stretch out beyond these potentially restrictive and artificial problems to apply their learning to situations similar to those they will meet as adults in their working and personal lives. Whereas the primary objective of many national examination systems is to ascertain whether or not the student has acquired the necessary mathematical foundation to see him, or her, into higher education in some specialism, the primary objective in PISA is to establish the extent to which those now leaving school will be able to meet the mathematical demands of living at the beginning of the 21st Century.

The tasks simulate, as closely as possible within the limitations of an assessment context, situations that students could well encounter in their present and future lives, with many items drawn from real-life examples provided by the participating countries. An expert group of mathematical educators was responsible for the selection, and formulation, of items. They applied themselves to ensuring that the final selection of items effectively tapped into students' ability to activate the mathematical knowledge and skills needed to solve such problems.

As problems in everyday life rarely present themselves with the mathematical route to their solution apparent, so in PISA, students have to decode the tasks set and translate them into a suitable mathematical form before they can start to solve them. Process and situation are key factors in the PISA concept of mathematical literacy. Simple technical, or even conceptual, competence in mathematics is not enough. Students must also demonstrate the ability to unravel the core of each task and, when a unique solution exists, find a suitable mathematical model for solving it or, when one does not exist, the model they judge provides the best answer to the task set.

How Mathematical Literacy was Assessed in PISA 2003

The assessment of mathematics was set in a framework defined by three factors[6]: the mathematical content of the tasks, the processes required for interpreting each task in a mathematical form, and the various situations in which such tasks might be encountered.

Four content zones were covered:

- space and shape
- change and relationships
- quantity, and
- uncertainty.

Tasks were categorised at three levels of process:

- the reproduction cluster
- the connection cluster, and
- the reflection cluster.

---

[6] See: {OECD, 2003. The PISA 2003 Assessment Framework.}

Tasks were set in four situational contexts:

personal
educational and occupational
public, and
scientific.

A brief description of these follows.

Four Content Areas

Space and shape, as the name implies, draws upon the discipline of geometry. It requires students to recognise similarities and differences in the shapes of objects when presented in different representations and in different dimensions, as well as the concepts of relative position and movement.

Change and relationships relates most closely to algebra. It involves, besides an understanding of the functional dependency between variables, an awareness of inequalities, equivalence, divisibility, etc, as well as a recognition that relationships can be expressed in various mathematical forms, and that changing between representations may be the key to solving a problem.

Quantity covers those aspects of mathematics bearing upon number. Students have to demonstrate competence in the many facets of this, from an understanding of relative size, and the use of numbers as representations of objective properties of bodies, through to the higher level of *quantitative reasoning*, which requires the understanding of the meaning of arithmetical operations.

Uncertainty lies within the area of probability and statistics. PISA regards the understanding of statistical ideas and the ability to follow statistical arguments as of increasing importance, if citizens are to participate effectively in democratically organised societies.

Process

The process of translating a problem to an appropriate mathematical formulation, is rarely achieved in one single step. A complete formulation often requires several levels of conceptual refinement. Similarly, the actual mathematics needed to solve it may require several different levels of proficiency. Some steps may be straightforward arithmetic, others may require careful algebraic manipulation. Finally, the complete answer to a problem may require several partial solutions, each at different conceptual levels. Therefore, real-life problems will rarely fall neatly into one single level of any framework devised to categorise process. Consequently, the main value of such a framework lies in confirming that the desired conceptual range is adequately covered by the collection of items used. It has less value for the reporting of results. The framework adopted by PISA assembled cognitive activities into three clusters, loosely hierarchical in structure. These are:

The reproduction cluster covers those competencies necessary to solve familiar, routine problems. Students essentially utilise practised knowledge, standard methods, and straightforward calculations.

The connections cluster comprises competencies that students have to deploy to solve those problems which, although set in a familiar situation, do not present an immediately recognisable solution. Such problems typically involve a greater degree of interpretation for their solution than those in the previous cluster.

The reflection cluster covers the highest levels of competencies required by PISA. Tasks in this cluster require some insight, reflection, and creativity. These tasks typically involve more mathematical elements than others and require students to explain and justify their reasoning and methods.

Situation

Students were presented with situations that they could conceivably meet in their own lives and which mathematical methods would help them analyse and resolve. Such situations fell into four broad categories: personal, educational/occupational, public, and scientific. A short description of each follows.

Personal situations relate directly to the student's own daily life. They have at their core the ways in which the individual perceives and is affected by an immediate, personal context. The student has to utilise their mathematics to appreciate, or interpret, some specific aspect of each situation.

Educational/occupational situations include settings that could arise in a student's school or work life. The core of these situations is how particular school and work settings present students with problems requiring a mathematical solution.

Public situations require students to observe aspects of their broader surroundings. They are generally situations located in the community and their core consists of the relationships that exist between the operative factors. A mathematical evaluation of the aspects of such relationships that have consequences for public life is wanted.

Scientific situations comprise those more abstract contexts typically involving a technological process, a theoretical one, or an explicitly mathematical problem. PISA includes within this situation abstract problems frequently confronted in the mathematics classroom, without attempt at contextualisation.

How the Mathematical Literacy Results are Reported

Results are presented in two principal ways.

The first uses a scale of scores obtained by modelling the patterns of item responses from each student. As each student sat just one booklet from the thirteen test booklets used in the assessment, statistical modelling of the responses is necessary to place all students on a common score scale. This scale was set to have a mean across OECD countries of 500 and a range such that two-thirds of students would score between 400 and 600[7]. Scales were derived for each of the four content areas, and for mathematics as a whole.

The second method of reporting results uses six 'proficiency levels' – descriptions of the kind of mathematical competency demonstrated by students. Summary descriptions of the proficiency levels are provided in Appendix B. Full descriptions can be found in the OECD publications arising from the study, which have been mentioned already. Descriptions of the highest and lowest levels are given below for convenience. Each test item used in PISA 2003 was matched to one of the six proficiency levels, and students were then placed at a specific proficiency level depending on how they had answered the set of items allocated to this level. More specifically, a student was placed at a particular proficiency level if he or she could be expected to answer correctly at least 50% of a hypothetical range of items spread evenly across the difficulty range for that level.

At Level 6, the highest level:

> *" .. students can conceptualise, generalise, and utilise information based on their investigations and modelling of complex problem situations. They can link different information sources and representations and flexibly translate among them. Students at this level are capable of advanced mathematical thinking and reasoning. These students can apply this insight and understanding along with a mastery of symbolic formal mathematical operations and relationships to develop new approaches and*

[7] The 2003 results were used for this. Thus, the scale is not the same as that in the report on PISA 2000.

strategies for attacking novel situations. Students at this level can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments, and the appropriateness of these to the original situations."

At Level 1:

"...students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are obvious and follow immediately from the given stimuli."

Summary of Mathematics Results for the OECD and for Scotland

Proficiency Levels

Results for the combined mathematics scale are given in Table 2.1. Figure 2.a presents the same results in graphical form and orders countries on the basis of the percentage of their students that reach Level 2 or above. Results for the individual broad content areas will be presented later.

Just below one-third (31%) of students across the OECD area as a whole performed at the top three levels of mathematics (Level 4 or above), and 3.5% at the highest level, Level 6. In broad terms, although OECD country results do vary widely, Belgium, Korea, and Japan, have the greatest proportions of their students achieving both the top three levels jointly, and Level 6, the topmost, as well. Mexico has the lowest proportion of its students at these levels.

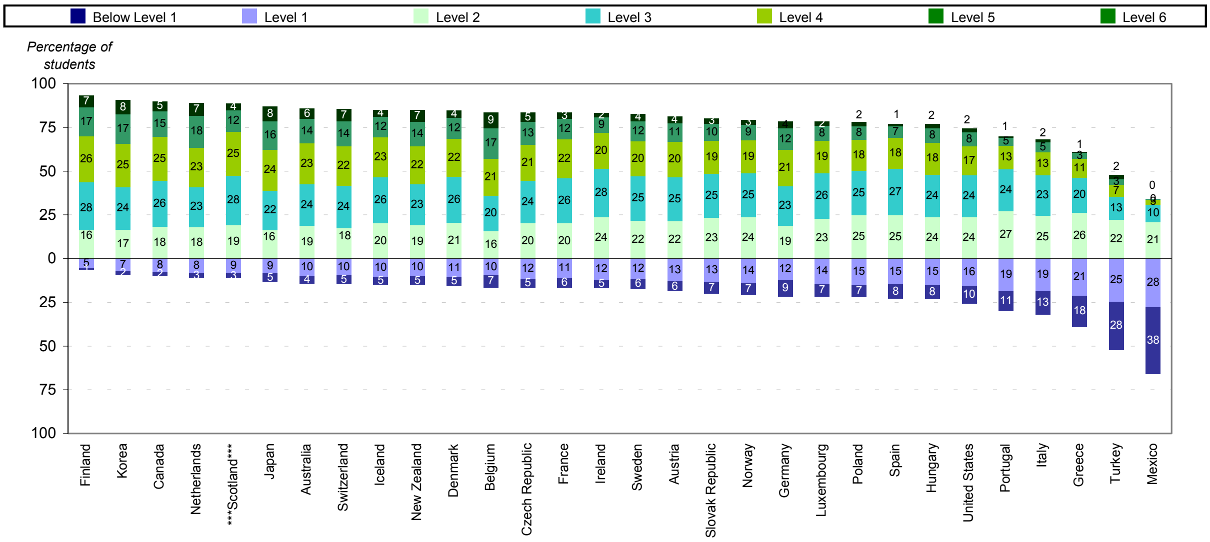
With 41.2% of Scottish students attaining Level 4 or better Scotland is well placed on the broader criterion covering Levels 4 to 6, but placed similarly to OECD students as a whole on the narrower criterion, with 3.9% of students achieving Level 6.

At the lower end of the scale, more students in Scotland, 89%, were operating at or above Level 2, compared with 74% in OECD countries as a whole. While Scotland thus appears well placed in respect of the poorer performing students, Level 2 merely requires that students can interpret and recognise situations in contexts that require no more than direct inference and that they can extract information from a single source and make use of a single representational mode. That 11% of our 15 year old students fail to reach Level 2, may give cause for concern.

Table 2.1. Percentage of students at each level of proficiency on the mathematics literacy scale

Country	Proficiency levels - mathematics													
	Below level 1 (below 335 score points)		Level 1 (from 336 to 420 score points)		Level 2 (from 421 to 482 score points)		Level 3 (from 483 to 544 score points)		Level 4 (from 545 to 606 score points)		Level 5 (from 607 to 669 score points)		Level 6 (above 670 score points)	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD Countries														
Australia	4.3	(0.4)	10.0	(0.5)	18.6	(0.6)	24.0	(0.7)	23.3	(0.6)	14.0	(0.5)	5.8	(0.4)
Austria	5.6	(0.7)	13.2	(0.8)	21.6	(0.9)	24.9	(1.1)	20.5	(0.8)	10.5	(0.9)	3.7	(0.5)
Belgium	7.2	(0.6)	9.3	(0.5)	15.9	(0.6)	20.1	(0.7)	21.0	(0.6)	17.5	(0.7)	9.0	(0.5)
Canada	2.4	(0.3)	7.7	(0.4)	18.3	(0.6)	26.2	(0.7)	25.1	(0.6)	14.8	(0.5)	5.5	(0.4)
Czech Republic	5.0	(0.7)	11.6	(0.9)	20.1	(1.0)	24.3	(0.9)	20.8	(0.9)	12.9	(0.8)	5.3	(0.5)
Denmark	4.7	(0.5)	10.7	(0.6)	20.6	(0.9)	26.2	(0.9)	21.9	(0.8)	11.8	(0.9)	4.1	(0.5)
Finland	1.5	(0.2)	5.3	(0.4)	16.0	(0.6)	27.7	(0.7)	26.1	(0.9)	16.7	(0.6)	6.7	(0.5)
France	5.6	(0.7)	11.0	(0.8)	20.2	(0.8)	25.9	(1.0)	22.1	(1.0)	11.6	(0.7)	3.5	(0.4)
Germany	9.2	(0.8)	12.4	(0.8)	19.0	(1.0)	22.6	(0.8)	20.6	(1.0)	12.2	(0.9)	4.1	(0.5)
Greece	17.8	(1.2)	21.2	(1.2)	26.3	(1.0)	20.2	(1.0)	10.6	(0.9)	3.4	(0.5)	0.6	(0.2)
Hungary	7.8	(0.8)	15.2	(0.8)	23.8	(1.0)	24.3	(0.9)	18.2	(0.9)	8.2	(0.7)	2.5	(0.4)
Iceland	4.5	(0.4)	10.5	(0.6)	20.2	(1.0)	26.1	(0.9)	23.2	(0.8)	11.7	(0.6)	3.7	(0.4)
Ireland	4.7	(0.6)	12.1	(0.8)	23.6	(0.8)	28.0	(0.8)	20.2	(1.1)	9.1	(0.8)	2.2	(0.3)
Italy	13.2	(1.2)	18.7	(0.9)	24.7	(1.0)	22.9	(0.8)	13.4	(0.7)	5.5	(0.4)	1.5	(0.2)
Japan	4.7	(0.7)	8.6	(0.7)	16.3	(0.8)	22.4	(1.0)	23.6	(1.2)	16.1	(1.0)	8.2	(1.1)
Korea	2.5	(0.3)	7.1	(0.7)	16.6	(0.8)	24.1	(1.0)	25.0	(1.1)	16.7	(0.8)	8.1	(0.9)
Luxembourg	7.4	(0.4)	14.3	(0.6)	22.9	(0.9)	25.9	(0.8)	18.7	(0.8)	8.5	(0.6)	2.4	(0.3)
Mexico	38.1	(1.7)	27.9	(1.0)	20.8	(0.9)	10.1	(0.8)	2.7	(0.4)	0.4	(0.1)	0.0	(0.0)
Netherlands	2.6	(0.7)	8.4	(0.9)	18.0	(1.1)	23.0	(1.1)	22.6	(1.3)	18.2	(1.1)	7.3	(0.6)
New Zealand	4.9	(0.4)	10.1	(0.6)	19.2	(0.7)	23.2	(0.9)	21.9	(0.8)	14.1	(0.6)	6.6	(0.4)
Norway	6.9	(0.5)	13.9	(0.8)	23.7	(1.2)	25.2	(1.0)	18.9	(1.0)	8.7	(0.6)	2.7	(0.3)
Poland	6.8	(0.6)	15.2	(0.8)	24.8	(0.7)	25.3	(0.9)	17.7	(0.9)	7.8	(0.5)	2.3	(0.3)
Portugal	11.3	(1.1)	18.8	(1.0)	27.1	(1.0)	24.0	(1.0)	13.4	(0.9)	4.6	(0.5)	0.8	(0.2)
Scotland	2.8	(0.5)	8.5	(0.7)	19.2	(1.1)	28.2	(1.4)	25.1	(1.1)	12.2	(0.8)	3.9	(0.4)
Slovak Republic	6.7	(0.8)	13.2	(0.9)	23.5	(0.9)	24.9	(1.1)	18.9	(0.8)	9.8	(0.7)	2.9	(0.4)
Spain	8.1	(0.7)	14.9	(0.9)	24.7	(0.8)	26.7	(1.0)	17.7	(0.6)	6.5	(0.6)	1.4	(0.2)
Sweden	5.6	(0.5)	11.7	(0.6)	21.7	(0.8)	25.5	(0.9)	19.8	(0.8)	11.6	(0.6)	4.1	(0.5)
Switzerland	4.9	(0.4)	9.6	(0.6)	17.5	(0.8)	24.3	(1.0)	22.5	(0.7)	14.2	(1.1)	7.0	(0.9)
Turkey	27.7	(2.0)	24.6	(1.3)	22.1	(1.1)	13.5	(1.3)	6.8	(1.0)	3.1	(0.8)	2.4	(1.0)
United States	10.2	(0.8)	15.5	(0.8)	23.9	(0.8)	23.8	(0.8)	16.6	(0.7)	8.0	(0.5)	2.0	(0.4)
OECD student mean	***11.0***	***(0.3)***	***14.6***	***(0.3)***	***21.2***	***(0.3)***	***22.4***	***(0.3)***	***17.6***	***(0.2)***	***9.6***	***(0.2)***	***3.5***	***(0.2)***
OECD country mean	***8.2***	***(0.2)***	***13.2***	***(0.2)***	***21.1***	***(0.1)***	***23.7***	***(0.2)***	***19.1***	***(0.2)***	***10.6***	***(0.1)***	***4.0***	***(0.1)***

Figure 2.a. Percentage of students at each level of proficiency on the combined mathematics scale



Source: OECD PISA 2003

Mean Scores in Mathematics Literacy

Table 2.2 gives the mean scores for the 29 OECD countries and Scotland, along with mean scores for male students and female students separately.

Table 2.2. Student performance on the mathematics scale, all students and by gender

Country	Mathematics scale score							
	All students		**Females**		**Males**		**Difference**	
	Mean Score	**S.E.**	**Mean Score**	**S.E.**	**Mean Score**	**S.E.**	**Score dif.**	**S.E.**
OECD Countries and Scotland								
Australia	524	(2.1)	522	(2.7)	527	(3.0)	5	(3.8)
Austria	506	(3.3)	502	(4.0)	509	(4.0)	8	(4.4)
Belgium	529	(2.3)	525	(3.2)	533	(3.4)	8	(4.8)
Canada	532	(1.8)	530	(1.9)	541	(2.1)	11	(2.1)
Czech Republic	516	(3.5)	509	(4.4)	524	(4.3)	15	(5.1)
Denmark	514	(2.7)	506	(3.0)	523	(3.4)	17	(3.2)
Finland	544	(1.9)	541	(2.1)	548	(2.5)	7	(2.7)
France	511	(2.5)	507	(2.9)	515	(3.6)	9	(4.2)
Germany	503	(3.3)	499	(3.9)	508	(4.0)	9	(4.4)
Greece	445	(3.9)	436	(3.8)	455	(4.8)	19	(3.6)
Hungary	490	(2.8)	486	(3.3)	494	(3.3)	8	(3.5)
Iceland	515	(1.4)	523	(2.2)	508	(2.3)	-15	(3.5)
Ireland	503	(2.4)	495	(3.4)	510	(3.0)	15	(4.2)
Italy	466	(3.1)	457	(3.8)	475	(4.6)	18	(5.9)
Japan	534	(4.0)	530	(4.0)	539	(5.8)	8	(5.9)
Korea	542	(3.2)	528	(5.3)	552	(4.4)	23	(6.8)
Luxembourg	493	(1.0)	485	(1.5)	502	(1.9)	17	(2.8)
Mexico	385	(3.6)	380	(4.1)	391	(4.3)	11	(3.9)
Netherlands	538	(3.1)	535	(3.5)	540	(4.1)	5	(4.3)
New Zealand	523	(2.3)	516	(3.2)	531	(2.8)	14	(3.9)
Norway	495	(2.4)	492	(2.9)	498	(2.8)	6	(3.2)
Poland	490	(2.5)	487	(2.9)	493	(3.0)	6	(3.1)
Portugal	466	(3.4)	460	(3.4)	472	(4.2)	12	(3.3)
Scotland	524	(2.3)	520	(2.9)	527	(3.3)	7	(4.1)
Slovak Republic	498	(3.3)	489	(3.6)	507	(3.9)	19	(3.7)
Spain	485	(2.4)	481	(2.2)	490	(3.4)	9	(3.0)
Sweden	509	(2.6)	506	(3.1)	512	(3.0)	7	(3.3)
Switzerland	527	(3.4)	518	(3.6)	535	(4.7)	17	(4.9)
Turkey	423	(6.7)	415	(6.7)	430	(7.9)	15	(6.2)
United States	483	(2.9)	480	(3.2)	486	(3.3)	6	(2.9)
OECD student mean	***489***	***(1.1)***	***484***	***(1.3)***	***494***	***(1.3)***	***10***	***(1.4)***
OECD country mean	***500***	***(0.6)***	***494***	***(0.8)***	***506***	***(0.8)***	***11***	***(0.8)***

Positive differences indicate that males perform better than females,
Negative differences indicate that females perform better than males.

As previously noted, the mean score for the mathematics scale was set at 500 for the OECD overall. Scotland's mean score of 524 is therefore significantly above the OECD average. Eight OECD countries have mean scores higher than Scotland. These are: Australia, Belgium, Canada, Finland, Korea, Japan, The Netherlands, and Switzerland. However, the statistical uncertainty associated with extrapolating these test scores to the average performance to be expected from all 15 year old students means that only for Finland, Korea, and The Netherlands can it definitely be said that their national attainment is better than Scotland's. For the other five this may be the case, but it is not definitely so. Nine OECD countries have mean scores that do not differ significantly from the Scottish one and 17 have mean scores that are significantly lower. Table 2.3 summarises the position.

Table 2.3: OECD countries whose mean scores differ significantly from the Scottish mean, or do not differ significantly from this.

Significantly higher mean score than Scotland (3 countries)	Mean score not significantly different from that for Scotland (9 countries)	Significantly lower mean score than Scotland (17 countries)	
Finland	Australia	Austria	Norway
Korea	Belgium	France	Poland
Netherlands	Canada	Germany	Portugal
	Czech Republic	Greece	Slovak Republic
	Denmark	Hungary	Spain
	Iceland	Ireland	Sweden
	Japan	Italy	Turkey
	New Zealand	Luxembourg	United States
	Switzerland	Mexico	

Changes in Mean Scores in Mathematics Between PISA 2000 and PISA 2003

As mathematics was the primary domain in PISA 2003 more mathematics items were used for this survey than in PISA 2000: 85 items in 2003, compared with 31 in PISA 2000. Twenty (20) of the PISA 2000 items (those not publicly released in the reports on that study) were re-used in 2003 to provide a common core for the two studies.

Several factors conspire against drawing strong conclusions about changes in countries' mean scores between the two surveys. The three principal factors are: firstly, the relatively small number of items linking the two studies; secondly, the fact that the items carried forward were not distributed evenly across the four content areas of mathematics, but came largely from just two, namely *Space and Shape*, and *Change and Relationships*; and thirdly, that context changes between the booklets used in 2000 and those in 2003 introduce appreciable statistical uncertainty into the matching of the scales for each survey. For these reasons, comparisons between the two surveys are necessarily general in nature and confined to the two content areas mentioned.

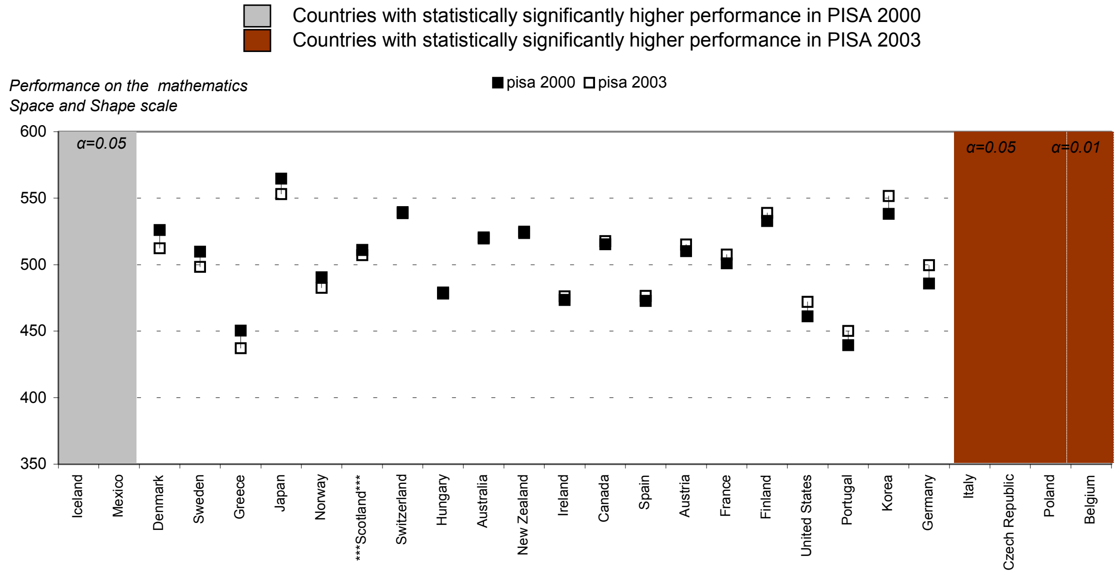
Space and Shape

Performance could be compared across 25 OECD countries and Scotland. Figure 2.b shows the comparison in graphical form. Small black squares denote the PISA 2000 mean scores and small white squares, the PISA 2003 ones. Shaded areas show countries with significantly different mean scores between the two surveys.

The average score across these 25 countries as a whole did not change significantly. On a common scale, the mean score in PISA 2000 was 494, while that in PISA 2003 was 496 points. Four OECD countries show a significant improvement in mean score and two a significant drop. The four showing these improved scores are: Belgium, The Czech Republic, Italy, and Poland. The two showing a drop in scores are: Iceland and Mexico.

Scotland shows a slight, though non-significant drop in performance, with a mean score in PISA 2000 of 511 points and in PISA 2003, of 507 points. This drop was essentially uniform at all levels of student attainment. Comparing scores in 2000 with those in 2003 for students at a number of levels of attainment, the 5th, 10th, 25th 75th, 90th and 95th percentile levels, ie the 5% of students at the lowest level through to the 5% at the highest level, scores change by 5 points at most. A difference of 5 points is well within the plausible limits of statistical uncertainty.

Figure 2.b. Differences in scores between PISA 2000 and PISA 2003 on the mathematics *space & shape* scale



Source: OECD PISA 2003

Change and Relationships

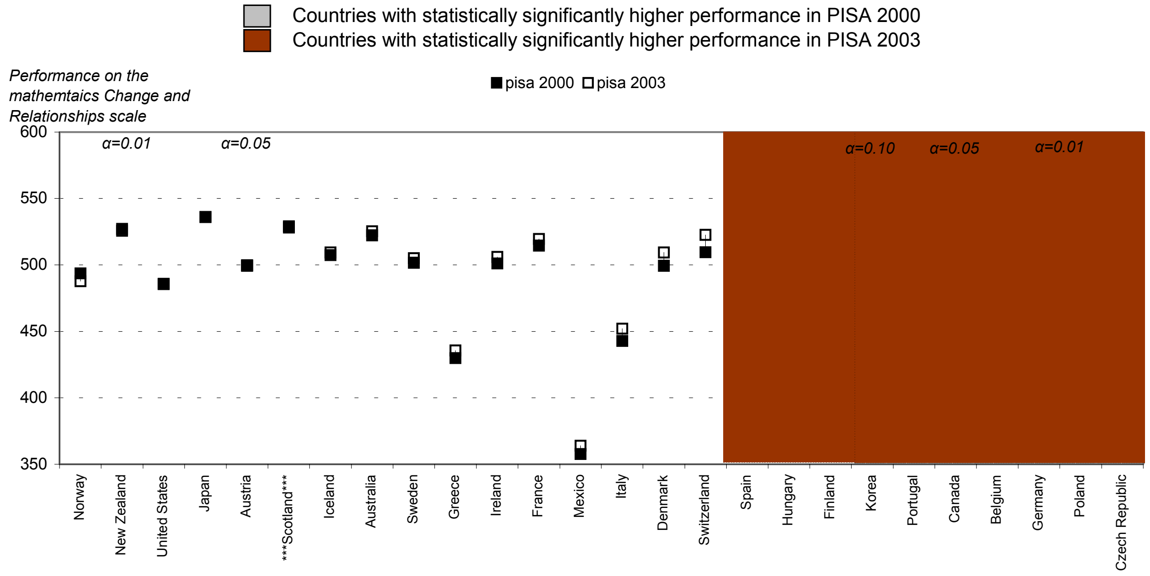
Performance in this content area could be compared also across 25 OECD countries and Scotland. Figure 2.c shows the comparison in graphical form. For these the average score did, in contrast to *Space and Shape*, show a significant change. On a common scale, the mean score in PISA 2000 was 488, while that in PISA 2003 was 499 points. Ten OECD countries show a significant improvement in mean score and none a significant drop. The ten showing improved scores are: Belgium, Canada, Czech Republic, Finland, Germany, Hungary, Korea, Poland, Portugal, and Spain.

The Scottish mean score decreased, but by just 1 point, between the two surveys. As with *Space and Shape*, there was no significant difference in performance observable at any level of attainment. The greatest difference at any level of attainment in *Change and Relationships* was just 3 points.

Although not visible in Figure 2.c, all but four OECD countries showed an increase in attainment in this content area between 2000 and 2003. This would imply that it is an area that many countries are focusing on developing, in which case Scotland's relative stand-still may be worrying. On the other hand, Scotland's score of 529 is already one of the highest and it may be that other countries are merely now catching up. Of the five countries with higher mean score than Scotland in PISA 2003[8]: two (Korea and Japan) already had higher scores than us in 2000; three (Belgium, Canada and Finland) made gains of between 14 and 22 points to enable them to leap-frog over Scotland.

[8] Although The Netherlands also had a higher mean score than Scotland in 'Change and Relationship' in PISA 2003, there was no data on its performance on this in PISA 2000 (The Netherlands failed to obtain the required sample size in PISA 2000).

Figure 2.c. Differences in scores between PISA 2000 and PISA 2003 on the mathematics *change & relationships* scale



Source: OECD PISA 2003

Gender Differences in Mathematical Literacy

Figure 2.d. Gender differences in mathematics mean score (score difference in favour of males)

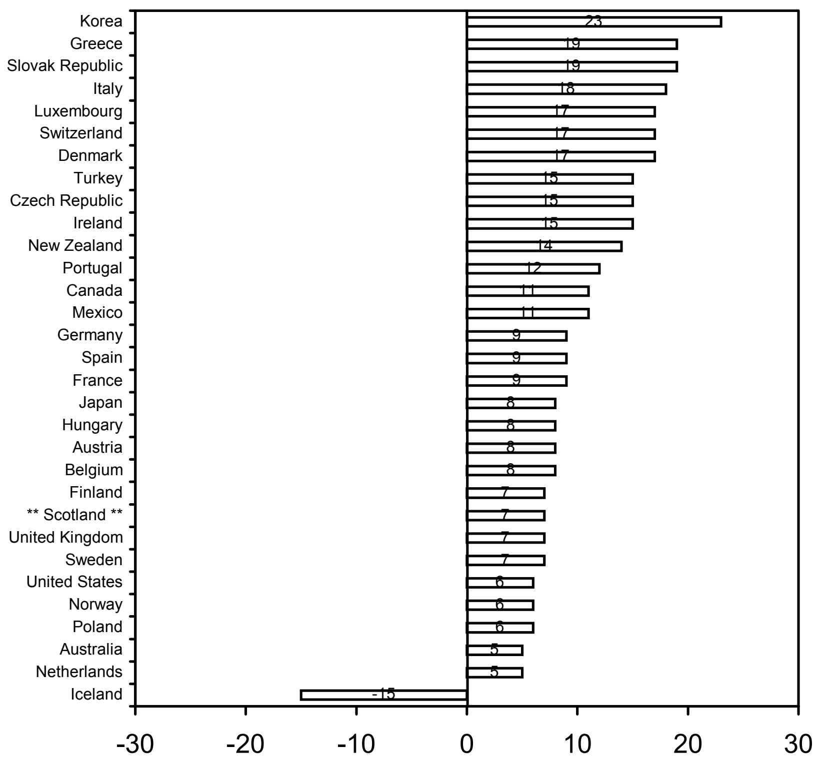
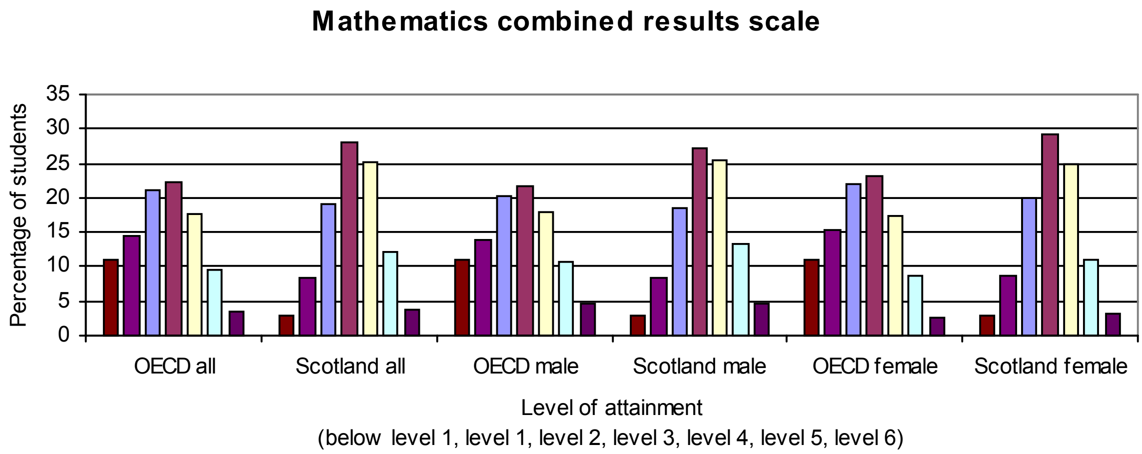


Figure 2.d shows the gender differences for the OECD countries and Scotland. This graph plots the difference between mean scores for male students and that for females. A bar to the right of the centre line means that male students scored higher than females, while one to the left means that female students did better than males. As can be seen, male 15 year

olds achieve higher than female in all the countries except Iceland. The range of score differences is large. Korea shows the highest difference, at 23 points but many countries also show large differences. For the 29 OECD countries as a whole, the average difference is close to 11 points. In Scotland the mean score for male students is 7 points above that for female students. In PISA 2000 Scottish male students were 5 points ahead of female. The almost universal better performance by males contrasts with the position in reading literacy where, as will be seen later, the situation reverses and females do better than males.

Figure 2.e. Mathematics combined results scale: comparison of male and female student attainment for OECD and Scotland



The final two paragraphs of this section focus on the performance of male and female students in Scotland relative to those in the OECD as a whole. The next section of the chapter details the results for each of the four content areas: *Space and Shape, Change and Relationships, Quantity,* and *Uncertainty.*

Figure 2.e shows the percentages of OECD and Scottish male and female students attaining the various levels of performance on the combined mathematics scale. Relative to the OECD, as already noted, a greater percentage of students in Scotland attain Levels 4, 5, and 6, the higher levels, though an almost equal percentage reaches Level 6. Equally, fewer Scottish students fall below Level 3 than in the OECD as a whole: while 69% of Scottish students attain Level 3 or better, the corresponding figure for the OECD as a whole is 53%. This pattern is mirrored, with slight variation, for both male and female students, with 71% of Scottish male students and 68% of female attaining Level 3 or better, compared with 55% and 52% respectively in the OECD overall.

Comparing students within Scotland, a greater percentage of male students than female attain the higher levels of performance, but this difference in attainment only shifts in their favour one level above Level 3, at Level 4. While 43% of male students attain Level 4 or above, just 39% of female students do so.

Mathematical Literacy in PISA 2003 – Attainment in the Four Broad Content Areas

This section presents the results for:

Space and shape
Change and relationships
Quantity, and
Uncertainty.

Table 2.4 gives the mean scores for each of these four content areas.

Table 2.4: Mean score and proportion of students reaching top three proficiency levels for each content area

	Mean score		Percentage of students gaining Level 3 or better	
	OECD country average	Scotland	Across the OECD area as a whole	Scotland
Space and shape	496	507	51	61
Change and relationships	499	529	54	70
Quantity	501	519	53	67
Uncertainty	502	536	54	74

Each content area shows a broadly similar pattern for Scotland and the whole OECD to that noted above for the overall performance levels. The mean scores for Scotland on every sub-scale are significantly higher than those for the OECD as a whole, and the proportions of Scottish students above Level 3 are, in general, substantially higher than those for the OECD as a whole.

The comparison of the Scottish mean scores with those for each of the individual 29 OECD countries is given in Table 2.5, and Table 2.6 summarises Scotland's position by giving the count of the number of OECD countries with mean scores; significantly above that for Scotland, and those not significantly different from, and significantly less than that for Scotland.

Table 2.5. Student performance on mathematics sub-scales, country mean scores

	Sh&Sp	Ch&Rel	Quant	Uncertainty
Australia	521	525	517	531
Austria	515	500	513	494
Belgium	530	535	530	526
Canada	518	537	528	542
Czech Republic	527	515	528	500
Denmark	512	509	516	516
Finland	539	543	549	545
France	508	520	507	506
Germany	500	507	514	493
Greece	437	436	446	458
Hungary	479	495	496	489
Iceland	504	509	513	528
Ireland	476	506	502	517
Italy	470	452	475	463
Japan	553	536	527	528
Korea	552	548	537	538
Luxembourg	488	487	501	492
Mexico	382	364	394	390
Netherlands	526	551	528	549
New Zealand	525	526	511	532
Norway	483	488	494	513
Poland	490	484	492	494
Portugal	450	468	465	471
***Scotland ***	507	529	519	536
Slovak Republic	505	494	513	476
Spain	476	481	492	489
Sweden	498	505	514	511
Switzerland	540	523	533	517
Turkey	417	423	413	443
United States	472	486	476	491
OECD student mean	**486**	**489**	**487**	**492**
OECD country mean	**496**	**499**	**501**	**502**

Table 2.6. Number of OECD countries with mean scores

	Number of OECD countries with mean scores:			
Sub-scale	Significantly below the Scottish mean	Not significantly different from the Scottish mean	Above the Scottish mean	Significantly above the Scottish mean
Space and shape	12	8	13	9
Change and relationships	19	8	6	2
Quantity	12	12	8	5
Uncertainty	20	8	4	1

Both in respect of the OECD taken as a whole and relative to the individual countries that constitute the OECD, Scottish students do very well in the content area of *Uncertainty*. The Scottish mean score is significantly above the OECD mean, just four countries have higher mean scores and only one of these has a significantly higher mean score than Scotland. *Change and Relationships* is the next strongest area. Six countries have higher mean scores than Scotland, two of them significantly higher. The Scottish mean score is significantly above the OECD mean score, and the percentage of students attaining Level 3 or above is well above the OECD percentage.

In the other two content areas, the picture is less clear. Scottish performance is good but not strong. In *Space and Shape*, the difference between the Scottish mean score and the OECD value, though significant, is the smallest of the four differences. Similarly the difference in percentages of students attaining Level 3 or above is the smallest of the four differences. Thirteen (13) countries have higher mean scores, nine significantly so. *Quantity* has the second smallest difference in mean score, and similarly the second smallest difference in percentage of students at Level 3 or above. Eight countries have higher mean scores than Scotland, five significantly so.

Parallel to what has been found for the combined mathematics scores, in each content area, male students attain higher mean scores than female both in the OECD as a whole and in Scotland. The percentages of male students attaining Level 3 or above are similarly greater than those for female students.

In *Space and Shape*, 63% of Scottish male students gained Level 3 or above compared with 58% of female students, while for the OECD as a whole, the corresponding comparison is 53% of males and 48% of females. The mean score for Scottish males (514) is 13 points higher than that for females (501), slightly less than the difference of 17 points in the OECD as a whole. Scottish males scored 9 points above OECD males: Scottish females, 13 points above OECD females[9].

In *Change and Relationships*, 72% of Scottish male students gained Level 3 or above, compared with 69% of female students. In the OECD as a whole, 55% of male students attained these levels, and 52% of female. The mean score for Scottish males (535) is 11 points higher than that for females (524), the same as the difference in the OECD as a

[9] Slight discrepancies between these four figures, and equally between those below, arise from rounding errors and imbalances in student counts between countries.

whole. Scottish males scored 30 points above OECD males; Scottish females, 31 points above OECD females.

In *Quantity*, Scottish male and female students did equally well in respect of proficiency Levels 3 and above, with 67% of both genders reaching these levels, while in the OECD as a whole, the two results were almost equal, at 54% and 52%, respectively. The mean score for Scottish males (521) was 4 points higher than that for females (517), slightly less than the difference of 6 in the OECD as a whole. Scottish males scored 17 points above OECD males; Scottish females, 19 points above OECD females.

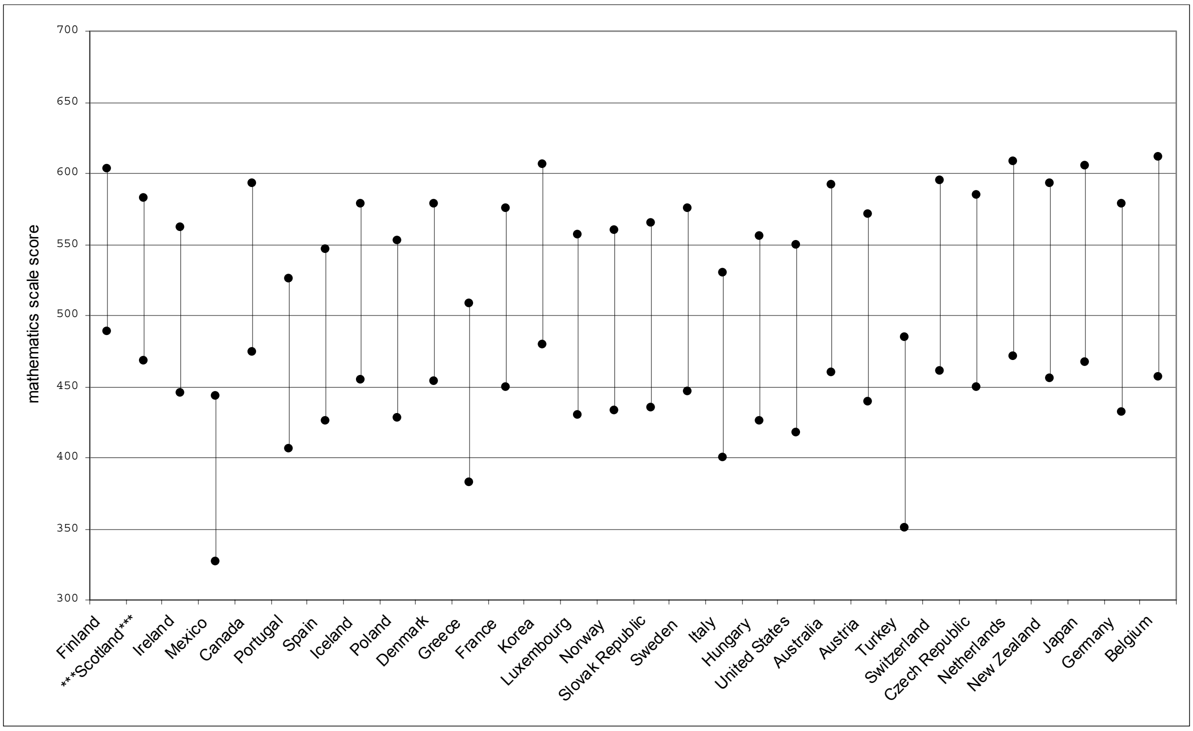
In *Uncertainty*, male and female students in Scotland performed almost equally well in respect of attainment at proficiency Levels 3 or above, with 75% of males and 74% of females reaching these levels, while in the OECD, the gap was wider, with 56% of male students attaining these levels and 52% of female. The mean score for Scottish males (538) was 5 points higher than that for females (533), markedly less than the difference of 13 points in the OECD as a whole. Scottish males scored 30 points above OECD males; Scottish females, 37 points above OECD females.

It is reassuring to note that Scotland has scored so highly in the area of *Uncertainty*. From 2001, SQA changed the arrangements for Standard Grade Mathematics examinations to introduce statistical content papers at all levels. (This was to maintain consistency with the Intermediate 1 and 2 courses which were introduced as part of the Higher Still initiative). In addition to all pupils being taught statistics as part of their National Qualification courses at S3 and S4, many schools introduced an element of statistics into their courses at S1 and S2 to improve the continuity and progression of pupils' learning.

Variation in Combined Mathematics Scores Between Low and High Achievers

Figure 2.f shows the score gap between students at the 75th percentile level, ie those at the top end of the attainment range, and those at the 25th level, ie, those at the bottom of the attainment range. Each pair of joined points corresponds to one of the 29 OECD countries or to Scotland. Countries are sequenced from left to right in increasing gap size (from smallest to largest difference). The smallest gap (115 points) is found in Finland, and the largest (155 points) in Belgium.

Figure 2.f. Difference in mathematics combined score between students at the 25th and 75th percentile levels of attainment



The graph compares the degree of equity in mathematics attainment between countries. That is, how great the achievement gap is between those who do well and those who do not. Is this gap narrow or wide? A country may have high overall mathematics achievement, as indicated by its mean score, but may provide a very unequal education to its students, or it may not. The gaps shown in this graph reveal where each country is placed in this respect.

Scotland is very well placed with a gap equal to that of Finland (115 points).

Chapter 3: Student Proficiency in Reading Literacy

How Reading Literacy is Defined

Reading literacy as defined in PISA focuses on the ability of students to use written information in situations that they encounter, or may encounter, in their lives. Reading literacy is 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. This notion goes beyond the traditional notion of decoding information and literal interpretation of what is written, towards more applied tasks.

How Reading Literacy was Assessed in PISA 2003

The assessment of reading for PISA 2003, as in PISA 2000, was framed along three dimensions: the type of reading task, the form and structure of the reading material, and the use for which the text was intended. Three broad types of task were used. Some required the retrieval of information, others the interpretation of the texts, and others called for reflection and evaluation of the texts. The text forms themselves were classified as 'narrative', 'expository', 'descriptive', 'discursive', 'instructive', and 'documentary'. Both continuous and non-continuous prose forms as well as charts, tables, diagrams, etc, were used. A wide variety of intended uses were covered under such categories as 'personal use', 'occupational use', and 'reading for education'. Further discussion of how the assessment was constructed is to be found in Chapter 6 of the International Report and in more detail in the PISA 2003 Assessment Framework (OECD, 2003).

Reading literacy was the main topic in PISA 2000, but a minor one in PISA 2003. Twenty eight of the 132 items used in PISA 2000, constituted the test of reading in PISA 2003. These 28 items were carefully chosen to give as balanced a picture of attainment in reading literacy as could be achieved within the limitations imposed by the study design, while at the same time providing a sound cross-section of the items used in PISA 2000. Each of the three aspects of reading: 'retrieval', 'interpretation', and 'reflection' were covered; as was the full difficulty range of the PISA 2000 items.

How the Reading Literacy Results are Reported

The more limited scope of the PISA 2003 assessment restricts the reporting of the 2003 results to the single combined scale of attainment, and not the three distinct sub-scales also used in PISA 2000[10]: 'retrieving information', 'interpreting texts', and 'reflecting and evaluating texts'. Results are, however, still reported by overall proficiency levels as well as by a scale score.

Appendix C gives a summary of the descriptions of performance expected at each proficiency level, and the full definitions can be found in the international reports on PISA 2000[11]. Descriptions of the lowest and highest proficiency levels are provided below for illustration. Each test item used in PISA 2003 was matched to one of the six proficiency levels and students were then placed at a specific proficiency level depending on how they had answered the set of items allocated to this level. More specifically, a student was placed at a particular proficiency level if he or she could be expected to answer correctly at least 50% of a hypothetical range of items spread evenly across the difficulty range for that level.

[10] A comprehensive treatment of performance in these three aspects of reading literacy can be found in the International report on that study.

[11] (Kirsch et al., 2003; OECD, 2001)

At Level 5, the highest level:

> *"Students proficient at Level 5 on the reading literacy scale are capable of completing sophisticated reading tasks, such as managing information that is difficult to find in unfamiliar texts; showing detailed understanding of such texts and inferring which information in the text is relevant to the task; and being able to evaluate critically and build hypotheses, draw on specialised knowledge, and accommodate concepts that may be contrary to expectations."*

At Level 1:

> *"Students at Level 1 can recognise an author's main theme or purpose in a text about a familiar topic, if this is prominent; they can make connexions between information in a text and common everyday knowledge; and they can locate one or more pieces of explicitly stated information in a text."*

The PISA 2003 results are scaled on the same scale as used in PISA 2000. This scale was set to a mean of 500 for the 27 OECD countries that participated in that study and a range such that two-thirds of students scored between 400 and 600. As the Slovak Republic and Turkey joined the OECD in 2003 and The Netherlands met all the technical standards in 2003, but not in 2000, and conversely for the UK, 29 OECD countries are now included in the PISA 2003 results. For these 29 countries, the overall OECD mean for reading literacy is 494, while the range remains unchanged.

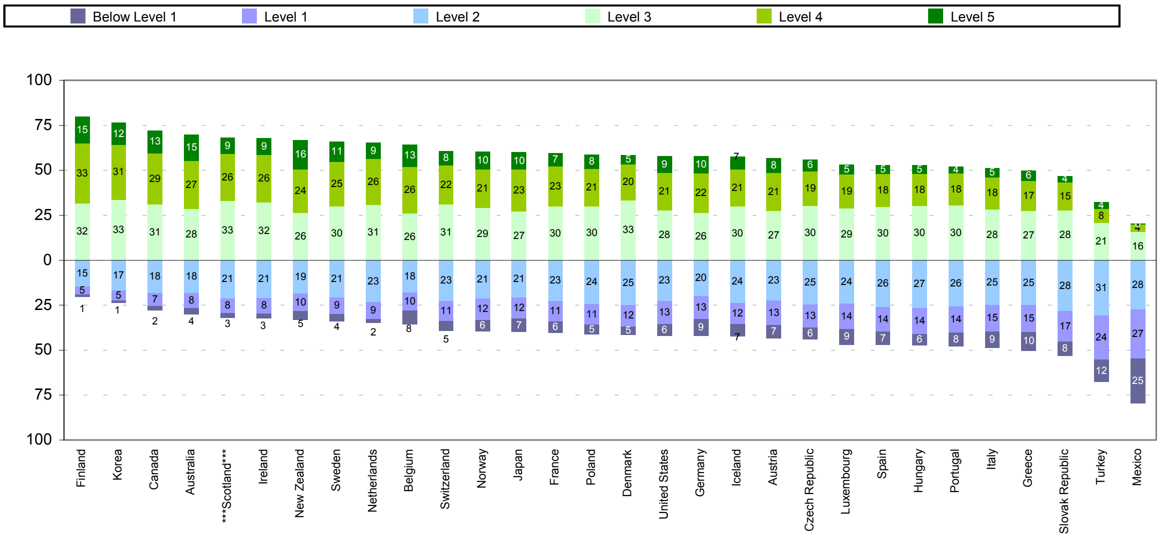
Summary of Reading Results for the OECD and Scotland.

Proficiency Levels

Table 3.1. Percentage of students at each level of proficiency on the reading literacy scale

Country	Proficiency levels											
	Below level 1 (less than 335 score points)		Level 1 (from 335 to 407 score points)		Level 2 (from 408 to 480 score points)		Level 3 (from 481 to 552 score points)		Level 4 (from 553 to 626 score points)		Level 5 (above 626 score points)	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD Countries and Scotland												
Australia	3.6	(0.4)	8.2	(0.4)	18.3	(0.6)	28.4	(0.8)	26.9	(0.8)	14.6	(0.7)
Austria	7.3	(0.8)	13.4	(1.0)	22.6	(1.0)	27.4	(1.0)	21.0	(1.0)	8.3	(0.8)
Belgium	7.8	(0.7)	10.0	(0.6)	18.2	(0.6)	26.0	(0.8)	25.4	(0.8)	12.5	(0.5)
Canada	2.3	(0.2)	7.3	(0.5)	18.3	(0.6)	31.0	(0.7)	28.6	(0.6)	12.6	(0.5)
Czech Republic	6.5	(0.9)	12.9	(0.9)	24.7	(1.0)	30.3	(1.3)	19.3	(1.1)	6.4	(0.6)
Denmark	4.6	(0.6)	11.9	(0.7)	24.9	(1.1)	33.4	(1.1)	20.0	(1.0)	5.2	(0.5)
Finland	1.1	(0.2)	4.6	(0.4)	14.6	(0.6)	31.7	(0.8)	33.4	(0.7)	14.7	(0.7)
France	6.3	(0.7)	11.2	(0.7)	22.8	(0.8)	29.7	(1.1)	22.5	(0.9)	7.4	(0.6)
Germany	9.3	(0.8)	13.0	(0.9)	19.8	(0.8)	26.3	(0.8)	21.9	(1.0)	9.6	(0.6)
Greece	10.2	(0.8)	15.0	(0.8)	25.0	(1.2)	27.3	(1.1)	16.8	(1.2)	5.7	(0.7)
Hungary	6.1	(0.7)	14.4	(0.9)	26.7	(0.9)	30.2	(1.1)	17.6	(1.1)	4.9	(0.6)
Iceland	6.7	(0.6)	11.8	(0.7)	23.9	(0.8)	29.7	(1.0)	20.9	(0.8)	7.1	(0.6)
Ireland	2.7	(0.5)	8.3	(0.7)	21.2	(1.2)	32.4	(1.3)	26.2	(1.2)	9.3	(0.7)
Italy	9.1	(0.9)	14.8	(0.8)	24.9	(0.8)	28.3	(1.0)	17.8	(0.7)	5.2	(0.3)
Japan	7.4	(0.8)	11.6	(0.8)	20.9	(1.0)	27.2	(1.1)	23.2	(1.1)	9.7	(0.9)
Korea	1.4	(0.3)	5.4	(0.6)	16.8	(1.0)	33.5	(1.2)	30.8	(1.1)	12.2	(1.1)
Luxembourg	8.7	(0.4)	14.0	(0.7)	24.2	(0.7)	28.7	(1.0)	19.1	(0.9)	5.2	(0.4)
Mexico	24.9	(1.5)	27.1	(1.2)	27.5	(1.0)	15.6	(1.0)	4.3	(0.6)	0.5	(0.1)
Netherlands	2.1	(0.5)	9.4	(0.9)	23.4	(1.1)	30.7	(1.3)	25.6	(1.1)	8.8	(0.7)
New Zealand	4.8	(0.5)	9.7	(0.6)	18.5	(0.9)	26.3	(0.9)	24.3	(0.9)	16.3	(0.8)
Norway	6.4	(0.6)	11.8	(0.8)	21.4	(1.2)	29.0	(1.0)	21.5	(0.8)	10.0	(0.7)
Poland	5.3	(0.5)	11.5	(0.7)	24.4	(0.8)	30.0	(0.9)	20.7	(0.9)	8.0	(0.6)
Portugal	7.6	(0.9)	14.4	(0.9)	25.9	(1.0)	30.5	(1.1)	17.9	(1.0)	3.8	(0.5)
Scotland	2.6	(0.4)	8.2	(0.8)	21.2	(1.0)	33.0	(1.2)	26.0	(1.0)	9.0	(0.7)
Slovak Republic	8.0	(0.8)	16.9	(1.0)	28.4	(1.0)	27.7	(1.1)	15.4	(0.7)	3.5	(0.4)
Spain	7.4	(0.7)	13.7	(0.7)	26.1	(0.7)	29.6	(0.8)	18.2	(0.9)	5.0	(0.5)
Sweden	3.9	(0.5)	9.4	(0.7)	20.7	(1.0)	29.9	(1.5)	24.8	(1.2)	11.4	(0.7)
Switzerland	5.4	(0.5)	11.3	(0.7)	22.7	(1.1)	30.9	(1.4)	21.9	(0.9)	7.9	(0.8)
Turkey	12.5	(1.2)	24.3	(1.5)	30.9	(1.4)	20.8	(1.4)	7.7	(1.1)	3.8	(1.2)
United States	6.5	(0.7)	12.9	(0.9)	22.7	(1.1)	27.8	(1.0)	20.8	(0.9)	9.3	(0.7)
OECD student percentages	***8.1***	***(0.3)***	***13.6***	***(0.3)***	***22.9***	***(0.4)***	***27.2***	***(0.4)***	***20.1***	***(0.3)***	***8.1***	***(0.2)***
OECD country percentages	***6.7***	***(0.1)***	***12.4***	***(0.2)***	***22.8***	***(0.2)***	***28.7***	***(0.2)***	***21.3***	***(0.2)***	***8.3***	***(0.1)***

Figure 3.a. Percentage of students proficient at each level of reading literacy



Source: OECD PISA 2003

Proficiency level results for reading literacy in the 29 OECD countries and Scotland are given in Table 3.1 and shown graphically in Figure 3.a.

In the whole OECD area, 8% of students are proficient at Level 5, the highest of the levels. More than 16% of students in New Zealand and more than 12% of the students in Australia, Belgium, Canada, Finland, and Korea are at this level. Nine percent (9%) of Scottish students reach this same level, slightly more than in the OECD as a whole. Less than 1% of students in Mexico reach this level.

If the results for the top three levels, Levels 3, 4 and 5, are combined, Australia, Canada, Finland, and Korea have over 70% of their students attaining these levels. Scotland has 68% of students attaining these levels. Overall, 55% of students in the OECD attain one of these three highest levels.

Taking a wider range, 78% of students in the OECD as a whole performed at Level 2 or above, but there are wide differences between countries. In Finland and Korea, about 94% of students reached these levels, but these are exceptions. In all other OECD countries, between 48 and 91% of students did so. Scotland is at the upper end of this range, with 89% of its students attaining Level 2 or above.

In the combined OECD area, 14% of students perform at Level 1, and 8% below Level 1, but there are wide differences between countries. In Finland and Korea, only around 5% of students perform at Level 1, and about 1% below it, but, again, these countries are exceptions. In all other OECD countries, between 10 and 52% of students perform at or below Level 1. One-third of the OECD countries have between 2 and 6% of students performing below Level 1. In Scotland, 8% of students performed at Level 1 and 3% below this.

Mean Scores in Reading Literacy

The first column of Table 3.2 gives the mean scores for the 29 OECD countries and Scotland, along with mean scores for male students and female students separately.

Table 3.2. Student performance on the reading scale, all students and by gender

Country	Reading scale							
	All students		Females		Males		Difference (M - F)1	
	Mean Score	S.E.	Mean Score	S.E.	Mean Score	S.E.	Score dif.	S.E.
OECD Countries and Scotland								
Australia	525	(2.1)	545	(2.6)	506	(2.8)	-39	(3.6)
Austria	491	(3.8)	514	(4.2)	467	(4.5)	-47	(5.2)
Belgium	507	(2.6)	526	(3.3)	489	(3.8)	-37	(5.1)
Canada	528	(1.7)	546	(1.8)	514	(2.0)	-32	(2.0)
Czech Republic	489	(3.5)	504	(4.4)	473	(4.1)	-31	(4.9)
Denmark	492	(2.8)	505	(3.0)	479	(3.3)	-25	(2.9)
Finland	543	(1.6)	565	(2.0)	521	(2.2)	-44	(2.7)
France	496	(2.7)	514	(3.2)	476	(3.8)	-38	(4.5)
Germany	491	(3.4)	513	(3.9)	471	(4.2)	-42	(4.6)
Greece	472	(4.1)	490	(4.0)	453	(5.1)	-37	(4.1)
Hungary	482	(2.5)	498	(3.0)	467	(3.2)	-31	(3.8)
Iceland	492	(1.6)	522	(2.2)	464	(2.3)	-58	(3.5)
Ireland	515	(2.6)	530	(3.7)	501	(3.3)	-29	(4.6)
Italy	476	(3.0)	495	(3.4)	455	(5.1)	-39	(6.0)
Japan	498	(3.9)	509	(4.1)	487	(5.5)	-22	(5.4)
Korea	534	(3.1)	547	(4.3)	525	(3.7)	-21	(5.6)
Luxembourg	479	(1.5)	496	(1.8)	463	(2.6)	-33	(3.4)
Mexico	400	(4.1)	410	(4.6)	389	(4.6)	-21	(4.4)
Netherlands	513	(2.9)	524	(3.2)	503	(3.7)	-21	(3.9)
New Zealand	522	(2.5)	535	(3.3)	508	(3.1)	-28	(4.4)
Norway	500	(2.8)	525	(3.4)	475	(3.4)	-49	(3.7)
Poland	497	(2.9)	516	(3.2)	477	(3.6)	-40	(3.7)
Portugal	478	(3.7)	495	(3.7)	459	(4.3)	-36	(3.3)
Scotland	516	(2.5)	527	(3.4)	504	(3.2)	-24	(4.4)
Slovak Republic	469	(3.1)	486	(3.3)	453	(3.8)	-33	(3.5)
Spain	481	(2.6)	500	(2.5)	461	(3.8)	-39	(3.9)
Sweden	514	(2.4)	533	(2.9)	496	(2.8)	-37	(3.2)
Switzerland	499	(3.3)	517	(3.1)	482	(4.4)	-35	(4.7)
Turkey	441	(5.8)	459	(6.1)	426	(6.8)	-33	(5.8)
United States	495	(3.2)	511	(3.5)	479	(3.7)	-32	(3.3)
OECD student mean	***488***	***(1.2)***	***503***	***(1.3)***	***472***	***(1.4)***	***-31***	***(1.4)***
OECD country mean	***494***	***(0.6)***	***511***	***(0.7)***	***477***	***(0.7)***	***-34***	***(0.8)***

Positive differences indicate that males perform better than females,
Negative differences indicate that females perform better than males.

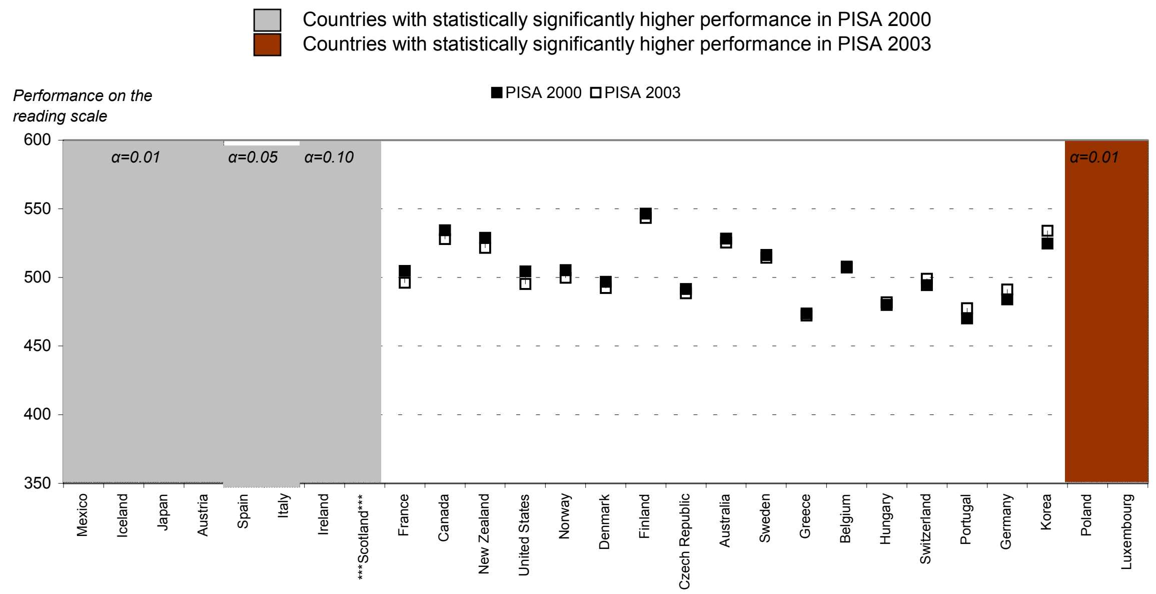
Only the three top performing OECD countries on the PISA 2003 reading literacy scale (Finland, and Korea and Canada) had mean reading scores that were significantly higher than that of Scotland. Scotland's mean score of 516 is (statistically) significantly above that for the OECD as a whole, and the table below lists the OECD countries whose mean scores are significantly greater or less than that for Scotland.

Table 3.3. OECD countries whose mean scores differ significantly from the Scottish mean.

Significantly higher mean score than Scotland (3 countries)	Mean score not significantly different from that for Scotland (6 countries)	Significantly lower mean score than Scotland (20 countries)		
Canada	Australia	Austria	Iceland	Portugal
Finland	Belgium	Czech Republic	Italy	Slovak Republic
Korea	Ireland	Denmark	Japan	Spain
	New Zealand	France	Luxembourg	Switzerland
	Sweden	Germany	Mexico	Turkey
	Netherlands	Greece	Norway	United States
		Hungary	Poland	

Changes in Mean Scores in Reading between PISA 2000 and PISA 2003

Figure 3.b. Differences in scores between PISA 2000 and PISA 2003 on the reading scale



Source: OECD PISA 2003

Comparative data is available for 26 of the 27 countries that participated in PISA 2000, the UK being the missing country, and for Scotland. Figure 3.b shows the two mean scores for these 27 countries sequenced from left to right in ascending order of the differences between the two scores. Countries on the left did better in PISA 2000, those on the right did better in PISA 2003.

None of the two mean scores for the 17 OECD countries lying between France and Korea differ significantly. Essentially attainment has not changed in these 17 countries. In two OECD countries, Luxembourg and Poland, mean performance rose significantly by 38 and

17 points respectively. In the other seven OECD countries, mean attainment fell significantly, by between 11 and 24 points.

Attainment in Scotland dropped by 11 points, a drop that is statistically significant, but at the 10% level of significance only. It is standard practice to measure statistical significance at the 5% level and at this level this is not a significant drop. However, as 11 points is one of the larger falls observed, it may be prudent not to ignore it, despite its borderline statistical position. While Scotland maintained its position just above the middle of the scale, average performance did shift closer to the OECD mean.

Scores in Reading for Students at Various Percentile Levels of Attainment

Improving attainment among the bottom 20% of students is an important policy priority for the Scottish Executive Education Department. The following table gives the scores attained by Scottish students at the lower levels of attainment[12] in PISA 2000 and PISA 2003. It shows the point below which the lowest performing 5%, 10% and 25% of students scored (i.e. the scores that mark off each of these percentiles).

Table 3.4: Scores attained by students at the lower end of the ability scale in reading literacy in PISA 2000 and PISA 2003

Percentile level	Score (in scale points) in PISA 2000	Score in PISA 2003	Difference (2003 – 2000) in scale points
5 %-ile	356	365	9
10 %-ile	394	403	9
25 %-ile	460	461	1

Scores for the two lowest groups have improved by approximately 9 scale points, but scores for the third group have not changed appreciably. An improvement of 9 points is not statistically significant.

In the OECD area as a whole, the performance of students at these three levels has fallen, not risen (by approximately 15 scale points at each of the percentiles).

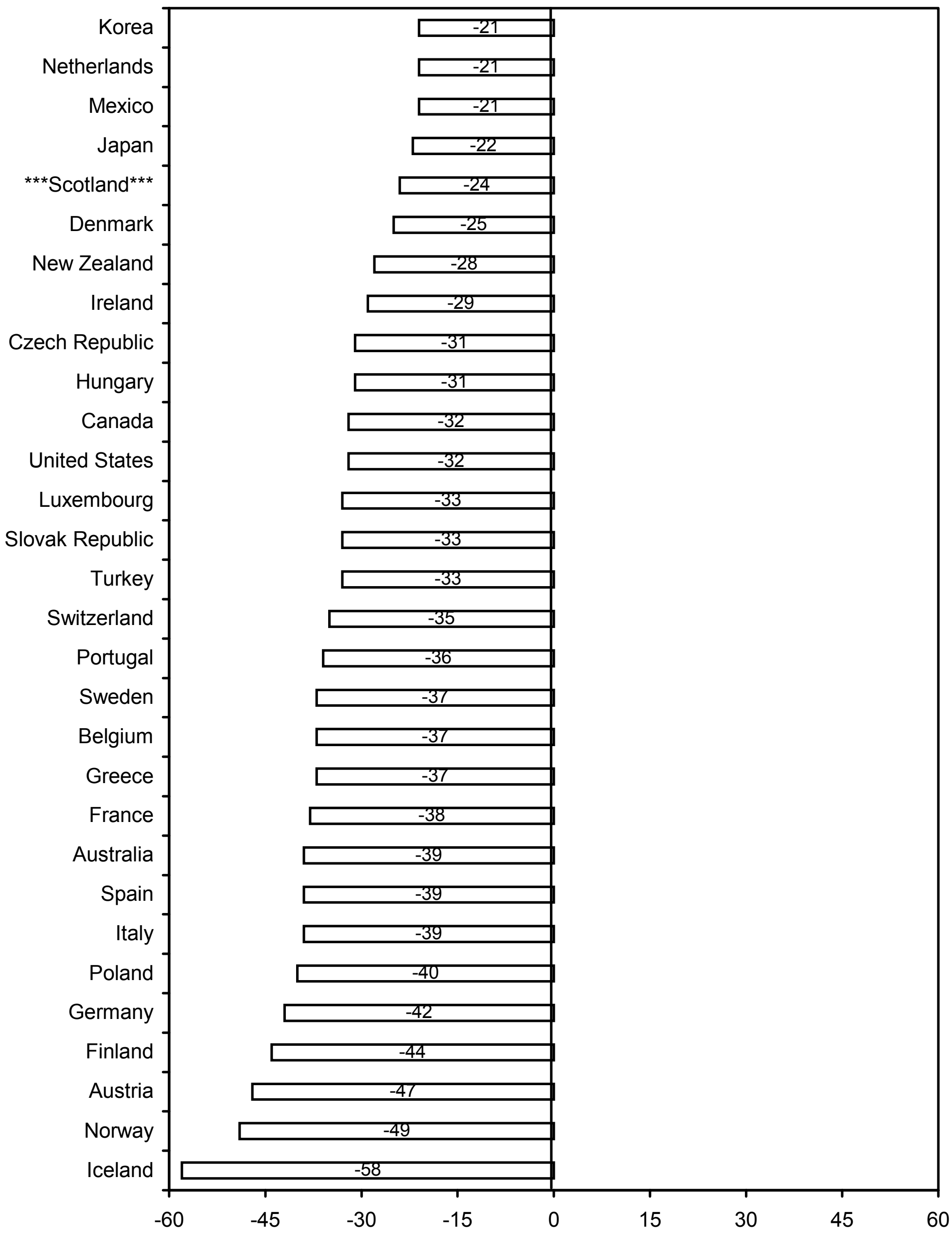
Overall, as just noted, the Scottish mean score in reading literacy fell by 11 points between PISA 2000 and PISA 2003. This overall decline is largely attributable to lower performance by pupils at the upper end of the ability range, those at the 75th, 90th and 95th %-ile, whose scores dropped by between 19 and 34 points. These differences are statistically significant. In the OECD as a whole, scores for these three groups also fell, but for each percentile by approximately 7 points only.

It is therefore reassuring that, although there has been a drop in Scotland's mean score, this drop has not been experienced across all ranges of ability. Given the emphasis on 'closing the gap' and promoting equity of attainment, it is particularly reassuring to note that, despite the trend across the OECD as a whole, the performance of our lowest 25% of students has improved (though not significantly).

[12] These results must be taken as indicative only and not as well established. The PISA 2000 results came from different tests to those used in PISA 2003, and it is uncertain how this alteration would affect pupils at the extremes of attainment in particular. While mean scores for whole countries would be little affected by the change, scores from students at the extremes may be more greatly so, particularly if the selection of items for PISA 2003 was unwittingly biased towards items that were either more difficult or more easy for these low attaining students.

Gender Differences in Reading

Figure 3.c. Gender differences in mean score in reading literacy (score difference in favour of males)



In every OECD country the mean reading score for female students is significantly higher than that for male students. The data are given numerically in Table 3.2 and shown graphically in Figure 3.c. This graph plots the difference between mean score for male students and that for females. A bar to the right of the centre line means male students scored higher than female, while one to the left means female students did better than males.

On average, across the OECD, the difference in performance between female and male students is 34 points, equivalent to half a proficiency level. There is, though, considerable variation between countries: from 58 points in Iceland, to 21 in Korea, The Netherlands and Mexico. In Scotland female students outperformed male students with a difference in mean score of 24 points. In PISA 2000, Scottish female students had a 30 point advantage over male students.

Variation in Reading Scores Between Low and High Achievers.

Figure 3.d. Difference in reading literacy scores between students at the 25th and 75th percentile levels of attainment

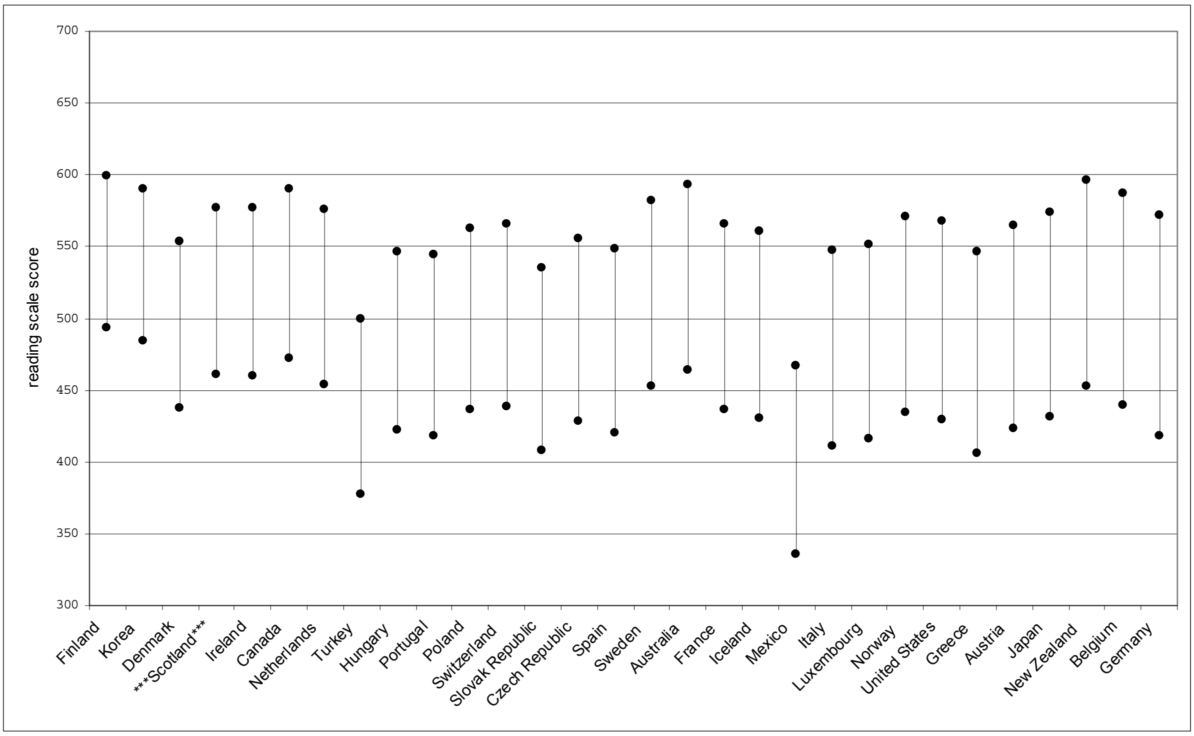


Figure 3.d shows the score gap between students at the 75th percentile level, ie those at the top end of the attainment range, and those at the 25th percentile level, at the bottom of the attainment range. Each pair of joined points corresponds to one of the 29 OECD countries or to Scotland. Countries are sequenced from left to right in increasing gap size. The smallest gap (105 points) is found in Finland, and the largest (153 points) in Germany.

The graph compares the degree of equity in reading attainment between countries. That is, how great the achievement gap is between those who do well and those who do not. Is this gap narrow or wide? A country may have high overall reading achievement, as indicated its mean score, but may provide a very unequal education to its students, or it may not. The gaps shown in this graph reveal where each country is placed in this respect.

Scotland is well placed in fourth position with a gap of 116 points, just slightly wider than Denmark's.

In PISA 2000, Scotland's score gap (between the 25th and 75th percentiles) in reading was 136 points. This means that between 2000 and 2003, our score gap narrowed by 15%. No other OECD country narrowed its score gap by as much as this.

Chapter 4: Student Proficiency in Scientific Literacy

How Scientific Literacy is Defined

As in the case of mathematics, the central concept defining science in PISA is its utility in the lives of all citizens in the early 21st Century, and not the focused requirements of students who are going forward to specialise in a scientific discipline. Understanding scientific data and reasoning are seen as increasingly relevant for effective participation in our present world. Consequently, the PISA items focus on assessing students' ability to: recognise scientific questions; identify what is involved in scientific investigations; relate scientific data to claims and conclusions; and communicate these aspects of science.

How Scientific Literacy was Assessed in PISA 2003

The science assessment comprised 35 items. These were the items used to assess scientific literacy in PISA 2000. No new items were added. The items were substantially as used in previously, although some were slightly modified in the light of experience.

A diverse range of items were employed. They varied in difficulty and covered a range of topics. The hardest items required complex conceptual skills, the less difficult required sound scientific thinking, and the easiest required straightforward recall and use of simple scientific knowledge.

In designing the assessment framework, the specialist science team took into account: scientific knowledge and concepts; scientific processes; and the science-based situations in which these need to be deployed. The framework covered three specific applications of science that raise issues for today's and tomorrow's citizens:

Science in life and health
Science in Earth and environment
Science in technology.

As there were relatively few scientific literacy test items, the framework domains were used only to ensure a due range of coverage in the item pool, and not for reporting purposes, and for the same reason, no proficiency levels were defined. Only scores are reported.

How the Scientific Literacy Results are Reported

The PISA 2003 results are scaled on the identical scale as used in PISA 2000. This scale was set to a mean of 500 for the 27 OECD countries that participated in that study and a range such that two-thirds of students scored between 400 and 600. As the Slovak Republic and Turkey joined the OECD in 2003 and The Netherlands met all the technical standards in 2003, but not in 2000, and conversely in the case of the UK, 29 countries are now included in the PISA 2003 results. For these 29 countries, the overall OECD mean for science literacy is 496 while the range remains unchanged.

Summary of Science Results for the OECD and Scotland

Mean Scores in Scientific Literacy

Table 4.1. Student performance on the science scale, all students and by gender

Country	Science scale							
	All students		Females		Males		Difference (M - F)1	
	Mean Score	S.E.	Mean Score	S.E.	Mean Score	S.E.	Score dif.	S.E.
OECD Countries and Scotland								
Australia	525	(2.1)	525	(2.8)	525	(2.9)	0	(3.8)
Austria	491	(3.4)	492	(4.2)	490	(4.3)	-3	(5.0)
Belgium	509	(2.5)	509	(3.5)	509	(3.6)	0	(5.0)
Canada	519	(2.0)	516	(2.2)	527	(2.3)	11	(2.6)
Czech Republic	523	(3.4)	520	(4.1)	526	(4.3)	6	(4.9)
Denmark	475	(3.0)	467	(3.2)	484	(3.6)	17	(3.2)
Finland	548	(1.9)	551	(2.2)	545	(2.6)	-6	(2.8)
France	511	(3.0)	511	(3.5)	511	(4.1)	0	(4.8)
Germany	502	(3.6)	500	(4.2)	506	(4.5)	6	(4.8)
Greece	481	(3.8)	475	(3.9)	487	(4.8)	12	(4.2)
Hungary	503	(2.8)	504	(3.3)	503	(3.3)	-1	(3.7)
Iceland	495	(1.5)	500	(2.4)	490	(2.4)	-10	(3.8)
Ireland	505	(2.7)	504	(3.9)	506	(3.1)	2	(4.5)
Italy	486	(3.1)	484	(3.6)	490	(5.2)	6	(6.3)
Japan	548	(4.1)	546	(4.1)	550	(6.0)	4	(6.0)
Korea	538	(3.5)	527	(5.5)	546	(4.7)	18	(7.0)
Luxembourg	483	(1.5)	477	(1.9)	489	(2.5)	13	(3.3)
Mexico	405	(3.5)	400	(4.2)	410	(3.9)	9	(4.1)
Netherlands	524	(3.1)	522	(3.6)	527	(4.2)	5	(4.7)
New Zealand	521	(2.4)	513	(3.4)	529	(3.0)	16	(4.2)
Norway	484	(2.9)	483	(3.3)	485	(3.5)	2	(3.6)
Poland	498	(2.9)	494	(3.4)	501	(3.2)	7	(3.3)
Portugal	468	(3.5)	465	(3.6)	471	(4.0)	6	(3.2)
Scotland	514	(2.7)	510	(4.0)	518	(3.7)	8	(5.5)
Slovak Republic	495	(3.7)	487	(3.9)	502	(4.3)	15	(3.7)
Spain	487	(2.6)	485	(2.6)	489	(3.9)	4	(3.9)
Sweden	506	(2.7)	504	(3.5)	509	(3.1)	5	(3.6)
Switzerland	513	(3.7)	508	(3.9)	518	(5.0)	10	(5.0)
Turkey	434	(5.9)	434	(6.4)	434	(6.7)	0	(5.8)
United States	491	(3.1)	489	(3.5)	494	(3.5)	5	(3.3)
OECD student mean	***496***	***(1.1)***	***493***	***(1.3)***	***499***	***(1.3)***	***6***	***(1.5)***
OECD country mean	***500***	***(0.6)***	***497***	***(0.8)***	***503***	***(0.7)***	***6***	***(0.9)***

Positive differences indicate that males perform better than females,
Negative differences indicate that females perform better than males.

The first column of Table 4.1 gives the mean science scores for each of the 29 OECD countries and Scotland.

The three top performing countries on the scientific literacy scale are Finland, Japan, and Korea. Scotland's score of 514 is significantly above the OECD average.

Extrapolating PISA results to the performance to be expected from all 15 year old students in each country introduces an element of statistical uncertainty into mean scores. Consequently, of the eight OECD countries with higher mean scores than Scotland, only the top three countries just mentioned can be said with certainty to have students who do better, on average, than those in Scotland. The higher mean tests scores in the other five countries may or may not imply better performance by their 15 year-olds. One cannot be certain. Similarly, of the 21 OECD countries with lower mean tests scores, only for 14 countries, those listed below, can it be said with certainty that their 15 year-old students do less well, on average, than Scottish 15 year-old students.

Table 4.2. OECD countries whose mean scores differ significantly from the Scottish mean

Significantly higher mean score than Scotland (3 countries)	Mean score not significantly different from that for Scotland (12 countries)		Significantly lower mean score than Scotland (14 countries)		
Finland	Australia	Hungary	Austria	Luxembourg	Slovak Republic
Japan	Belgium	Ireland	Denmark	Mexico	Spain
Korea	Canada	Netherlands	Greece	Norway	Turkey
	Czech Republic	New Zealand	Iceland	Poland	United States
	France	Sweden	Italy	Portugal	
	Germany	Switzerland			

Changes in Mean Scores in Science Between PISA 2000 and PISA 2003

As noted, essentially the same item set of 35 items was used in PISA 2003 as in PISA 2000, and the scoring scale was the same.

With relatively few items to cover the whole domain of science literacy, it would be unwise to place too great an emphasis on changes in countries' mean scores between the two surveys. However, the main trends observed are as follows.

Comparative data is available for 26 of the 27 OECD countries that participated in PISA 2000, the UK being the one missing country, and for Scotland. Figure 4.a shows the two mean scores for these 27 countries in ascending order, from left to right, of the difference between the two scores. The PISA 2000 scores are shown by small black squares; the 2003 ones by small open white square. Countries on the left did better in PISA 2000; those on the right did better in PISA 2003.

None of the 11 OECD countries lying between Ireland and Portugal, show a significant change in performance between 2000 and 2003. In the 10 OECD countries to the right of Portugal on the graph, mean performance rose by between 8 points (in Italy) and 40 points (in Luxembourg). In the other five OECD countries, mean attainment fell significantly, by between 10 points (in Canada[13]) and 30 points (in Austria).

Performance fell in Scotland between PISA 2000 and PISA 2003, but the 8 point drop, from 522 to 514 points, is not statistically significant. Scotland maintained its position just above the middle of the attainment scale.

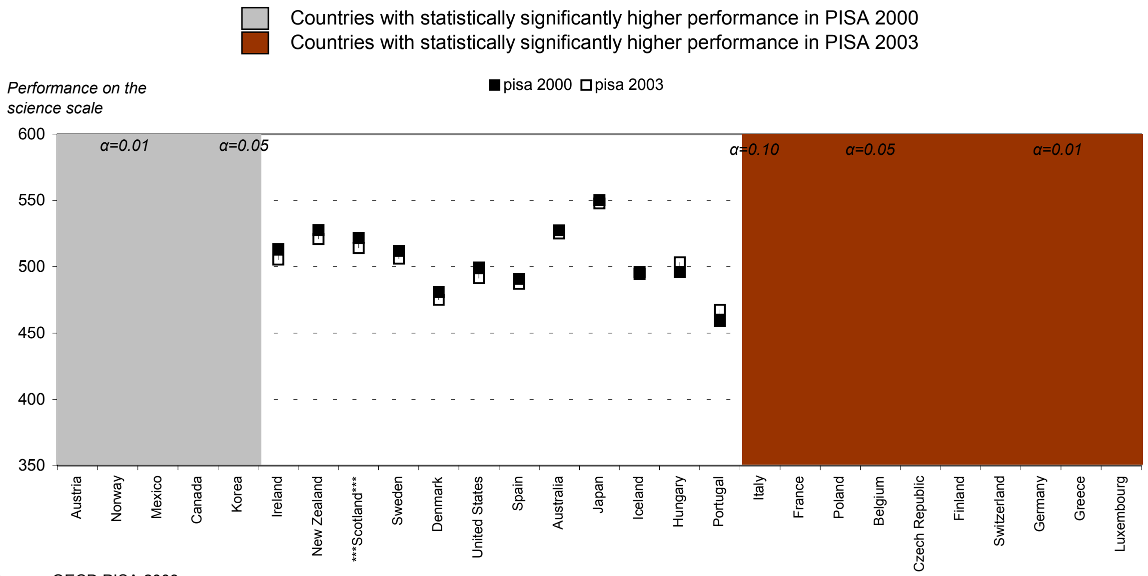
[13] Countries are actually sorted by the statistical significance of the differences in scores, not by the differences themselves, hence Canada appears to the left of Austria despite a smaller score difference.

Scores in Science for Students at Various Percentile Levels of Attainment

The drop in science literacy performance by students in Scotland between PISA 2000 and PISA 2003 was more or less uniform across the whole ability range. Performance by students at the very lowest levels of ability, those at the 5th %-ile and 10th %-ile levels fell by approximately 5 points, while those at the 25th 75th 90th and 95th %-ile levels fell by between 7 and 9 points.

In the OECD as a whole, students at the 5th, 10th, and 25th %-ile levels of the ability range dropped in score by between 12 and 16 points. Performance by those at the 75th level changed very little and performance by those at the 90th and 95th levels improved by 5 and 8 points respectively.

Figure 4.a. Differences in scores between PISA 2000 and PISA 2003 on the science scale



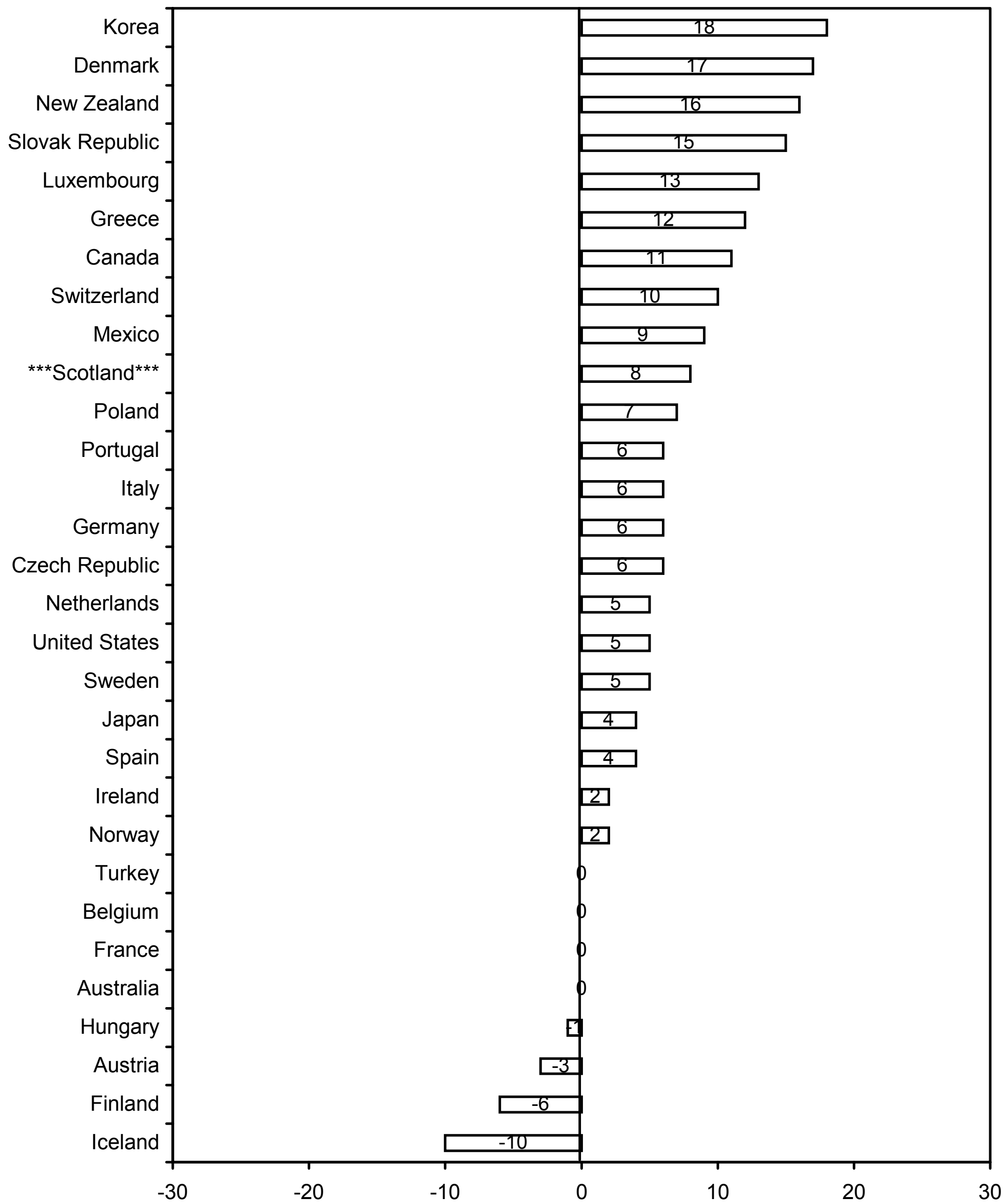
Source: OECD PISA 2003

Table 4.1 also gives the mean scores for male and female students, and the differences between these two are shown graphically in Figure 4.b. This graph plots the difference between mean score for male students and that for females. A bar to the right of the centre line means male students scored higher than female, while one to the left means female students did better than males.

In five OECD countries the mean scores for male and female students are equal to within 1 point. In 21 OECD countries the mean score for male students is higher than that for female students by between 1 and a maximum of 18 points (but the difference is statistically significant in only 11 of these countries – Korea through to Portugal). In three countries female students score higher than male students by between 2 and a maximum of 10 points (the difference is significant in Finland and Iceland). Korea shows the largest difference in favour of males, while Iceland, shows the largest difference in favour of females. Thus, while in general male students did better than female in the OECD at large, this is not universally the case.

Gender Differences in Science Literacy

Figure 4.b. Gender differences in mean score in science literacy (score difference in favour of males)



In Scotland there is a difference of 8 points in favour of male students, but this difference is not statistically significant. In PISA 2000[14], the gender difference in Scotland was just 1 point in favour of males. Quite possibly, therefore, Scotland's relatively high position in Figure 4.b may simply reflect the statistical uncertainty inherent in the assessment and not be a cause for concern.

[14] (Education and Young People Research Unit, 2002: page 10)

Variation in Science Scores Between Low and High Achievers

Figure 4.3. Difference in science literacy scores between students at the 25th and 75th percentile levels of attainment

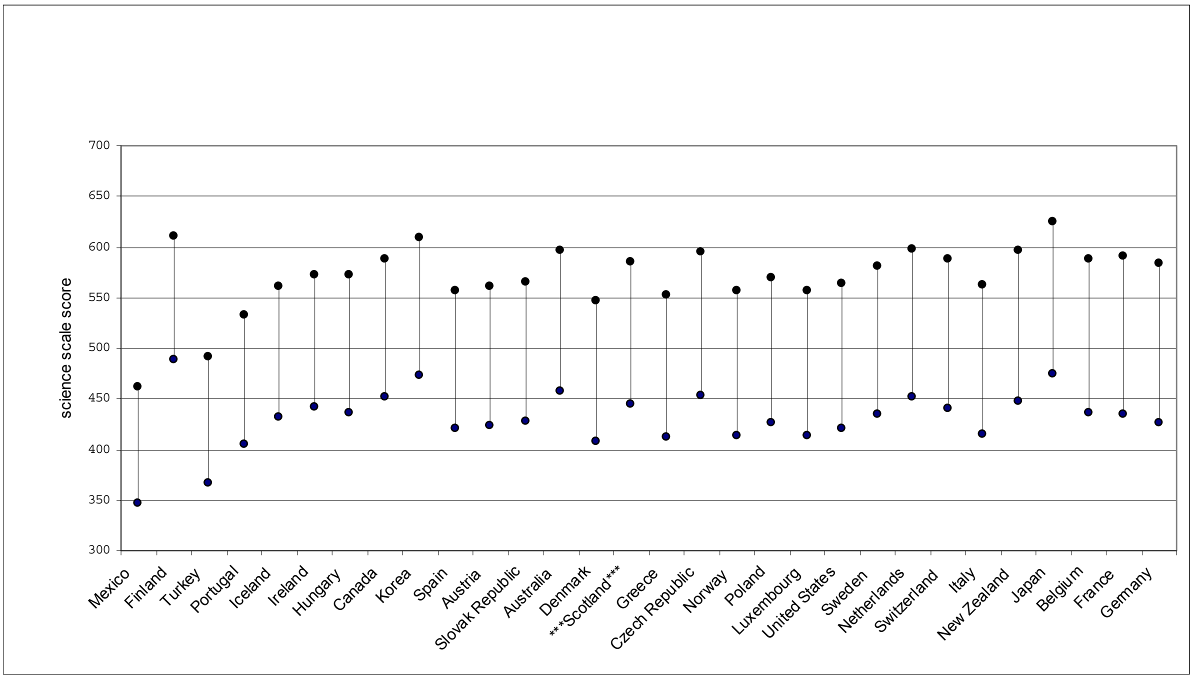


Figure 4.c shows the score gap between students at the 75th percentile level, i.e. those towards the top end of the attainment range, and those at the 25th level, i.e. those towards the bottom of the attainment range. Each pair of joined points corresponds to one of the 29 OECD countries or to Scotland. Countries are sequenced from left to right in increasing gap size. The smallest gap, of 115 points is found in Mexico, and the largest, of 157 points, in Germany.

This graph compares the degree of equity in science attainment between countries. That is, how great the achievement gap is between those who do well and those who do not. Is this gap narrow or wide? A country may have high overall science achievement, as indicated by its mean score, but may provide a very unequal education to its students; or it may not.

Scotland stands in the centre of the range, with a gap of 140 points. Scotland's rating on this equity scale is neither particularly good nor particularly bad, compared with the OECD countries. However, both in respect of quality, as measured by the overall mean score, and equity, as measured by the differences above, Scotland does fall well below what Finland shows can be achieved.

In PISA 2000, Scotland's score gap between the 25th and 75th percentiles for science was 142. This means that between 2000 and 2003 we narrowed the gap by only 2 score points. However, only three OECD countries managed to narrow their gap in science between these two sweeps of PISA: Belgium (by 1 score point); Denmark (by 3 score points) and Hungary (by 8 score points). In all other OECD countries the gap increased.

Appendix A

Results for the United Kingdom.

The England PISA 2003 sample fell short of the pre-agreed school and student level response rates. According to the technical standards established by OECD Member countries for PISA, this meant that for England (and therefore the United Kingdom) data to be fully included in the PISA report it was necessary to demonstrate that the England sample was free from bias.

Analysis of the England sample detected significant bias at student level. OECD analysts have concluded that at present the uncertainties surrounding the England sample are such that it is not possible to make reliable comparisons between the performance of the United Kingdom and that of other countries.

Annex 3 of the PISA 2003 international report contains the following note explaining how the United Kingdom did not meet in full the PISA technical standards and how subsequent bias analysis found that although there was no evidence of significant bias at the school level there was potential bias at the student level.

Extract from Annex 3 of the international report

In order to ensure that PISA yields reliable and internationally comparable data, OECD Member countries have established a process for the validation of all national data submissions. As the basis for this process, PISA established technical standards for the quality of datasets which countries must meet in order to be reported in OECD publications. These standards are described in detail in the PISA 2003 Technical Report. One of the requirements is that initial response rates should be 85% at the school level and 80% at the student level.

The United Kingdom fell significantly short of these standards, with an achieved initial response rate of 63% at the school level and 78% at the student level. The Technical Standards include an approved procedure through which countries with an initial school-level response rate of at least 65% could improve response rates through the use of designated replacement schools. For the United Kingdom, a school-level response rate of 95% was required but only 77% was achieved.

The results of a subsequent bias analysis provided no evidence for any significant bias of school-level performance results but did suggest that there was potential non-response bias at student levels. The PISA Consortium concluded that it was not possible to assess reliably the magnitude, or even the direction, of this non-response bias and to correct for this. As a result, it is not possible to say with confidence that the United Kingdom's sample results reliably reflect those for the national population, with the level of accuracy required by PISA. The mean performance of the responding sample of United Kingdom students was 508, 507 and 518 in mathematics, reading and science *respectively*. In the mathematics subscales the mean performance was 496 on the space and shape scale, 513 on the change and relationships scale, 520 on the uncertainty scale and 499 on the quantity scale. If negligible to moderate levels of bias are assumed, the United Kingdom mean performance would lie between 492 and 524 on the mathematical literacy scale, between 491 and 523 on the reading literacy scale, and between 502 and 534 on scientific literacy scale. The uncertainties surrounding the sample and its bias are such that scores for the United Kingdom cannot reliably be compared with those of other countries. They can also not be compared with the performance scores for the United Kingdom from PISA 2000.

The results are, however, accurate for within-country comparisons between subgroups (e.g. males and females) and for relational analyses. The results for the United Kingdom have, therefore, been included in a separate category below the results for the other participating countries. (And) In the main body of the report, where countries significantly above the

OECD average in a given indicator are mentioned, the United Kingdom will be included where relevant.

It should be noted that Scotland and Northern Ireland carried out an independent sample that met the PISA technical standards. Results for Scotland and Northern Ireland are reported in Annex B2 and are fully comparable with results from other OECD countries and with results from PISA 2000.

The material not in the OECD main report will be set out in a table and will appear on the OECD PISA website, to enable researchers to reproduce the international means and models.

Appendix B

Summary Proficiency Level Descriptors for Mathematical Literacy

Proficiency at Level 6

Students at Level 6 can conceptualise, generalise, and utilise information based on their investigations and modelling of complex problem situations. They can link different information sources and representations and flexibly translate among them. Students at this level are capable of advanced mathematical thinking and reasoning. These students can apply this insight and understandings along with a mastery of symbolic and formal mathematical operations and relationships to develop new approaches and strategies for attacking novel situations. Student at this level can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments, and the appropriateness of these to the original situations.

Proficiency at Level 5

Students at Level 5 can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare, and evaluate appropriate problem solving strategies for dealing with complex problems related to these models. Students at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insight pertaining to these situations. They can reflect on their actions and formulate and communicate their interpretations and reasoning.

Proficiency at Level 4

Students at Level 4 can work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic, linking them directly to aspects of real-world situations. Students at this level can utilise well-developed skills and reason flexibly, with some insight, in these contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments, and actions.

Proficiency at Level 3

Students at Level 3 can execute clearly described procedures, including those that require sequential decisions. They can select and apply simple problem solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They can develop short communications reporting their interpretations, results and reasoning.

Proficiency at Level 2

Students at Level 2 can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Students at this level can employ basic algorithms, formulae, procedures, or conventions. They are capable of direct reasoning and making literal interpretations of the results.

Proficiency at Level 1

Students at Level 1 can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are obvious and follow immediately from the given stimuli.

Appendix C

Summary Proficiency Level Descriptors For Reading Literacy

Proficiency at Level 5

Students proficient at Level 5 on the reading literacy scale are capable of completing sophisticated reading tasks, such as managing information that is difficult to find in unfamiliar texts; showing detailed understanding of such texts and inferring which information in the text is relevant to the task; and being able to evaluate critically and build hypotheses, draw on specialised knowledge, and accommodate concepts that may be contrary to expectations.

Proficiency at Level 4

Students proficient at Level 4 on the reading literacy scale are capable of difficult reading tasks, such as locating embedded information, construing meaning from nuances of language and critically evaluating a text.

Proficiency at Level 3

Students proficient at Level 3 on the reading literacy scale are capable of reading tasks of moderate complexity, such as locating multiple pieces of information, making links between different parts of a text, and relating it to familiar everyday knowledge.

Proficiency at Level 2

Students proficient at Level 2 are capable of basic reading tasks, such as locating straightforward information, making low-level inferences of various types, working out what a well-defined part of a text means, and using some outside knowledge to understand it.

Proficiency at Level 1

Reading literacy, as defined in PISA, focuses on the knowledge and skills required to apply “reading for learning” rather than on the technical skills acquired in “learning to read”. Since comparatively few young adults in OECD countries have not acquired technical reading skills, PISA does not seek to measure such things as the extent to which 15-year-old students are fluent readers or how well they spell or recognise words, but focuses on measuring the extent to which individuals are able to construct, expand and reflect on the meaning of what they have read in a wide range of texts common both within and beyond school. The simplest reading tasks that can still be associated with this notion of reading literacy are those at Level 1. Students proficient at this level are capable of completing only the least complex reading tasks developed for PISA, such as locating a single piece of information, identifying the main theme of a text or making a simple connection with everyday knowledge.

Proficiency below Level 1

Students performing below Level 1, are unlikely to demonstrate success on the most basic type of reading that PISA seeks to measure. This does not mean that they have no literacy skills but such students have serious difficulties in using reading literacy as an effective tool to advance and extend their knowledge and skills in other areas. Students with literacy skills below Level 1 may, therefore, be at risk not only of difficulties in their initial transition from education to work but also of failure to benefit from further education and learning opportunities throughout life.

References

Artelt, C., Baumert, J., Julius-McElvany, N., & Peschar, J. (2003). *Learners for Life: Student Approaches to Learning: Results form PISA 2000*: Organisation for Economic Co-operation and Development (Paris).

Cosgrove, J., Sofroniu, N., Kelly, A., & Shiel, G. (2003). A Teacher's Guide to the Reading Literacy Achievements of Irish 15 Year Olds: Outcomes of the OECD's Programme for International Student Assessment (PISA). Dublin: Educational Research Centre, St Patrick's College.

Department for Education and Skills. (2003). *The future of higher education (Cm 5375)*. The Stationery Office: HMSO.

Education and Young People Research Unit. (2002). *Programme for International Student Assessment: Scottish Report*. Edinburgh: Scottish Executive Education Department.

Education Statistics Publications. (2002). *Programme for International Student Assessment - Scotland Analysis*. Edinburgh: Scottish Executive.

Gill, B., Dunn, M., & Goddard, E. (2002). Student achievement in England: results in reading, mathematical and scientific literacy among 15-year olds from OECD PISA 2000 study. London: The Stationery Office.

Kirsch, I., de Jong, J., Lafontaine, D., McQueen, J., Mendelovits, J., & Monseur, C. (2003). *Reading for Change: Performance and Engagement across Countries Results from PISA 2000*: Organisation for Economic Co-operation and Development (Paris).

Lie Svein, L. P., Roe Astrid. (2003). *Northern Lights on PISA: Unity and Diversity in the Nordic Countries in PISA 2000*. OSLO: Department of Teacher Education and School Development, University of Oslo.

MLWIN. (2004). Centre for Multilevel Modelling.

OECD. (2001). Knowledge and Skills for Life: First Results from the OECD Programme for International Student Assessment (PISA) 2000: Organisation for Economic Co-operation and Development (Paris).

OECD. (2002). *Manual for the PISA 2000 Database*: Organisation for Economic Co-operation and Development (Paris).

OECD. (2002). Sample Tasks from the PISA 2000 Assessment: Reading, Mathematical and Scientific Literacy: Organisation for Economic Co-operation and Development (Paris).

OECD. (2002). *PISA 2000 Technical Report*: Organisation for Economic Co-operation and Development (Paris).

OECD. (2003). The PISA 2003 Assessment Framework: Mathematics, reading, science and problem solving knowledge and skills: Organisation for Economic Co-operation and Development (Paris).

OECD, & UNESCO-UIS. (2003). *Literacy Skills for the World of Tomorrow: Further results from PISA 2000*. Paris: Organisation for Economic Co-operation and Development

United Nations Educational, Scientific and Cultural Organization Institute for Statistics.

PISA, O. (2003). The OECD PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT.

Valijarvi, j., Linnakyla, P., Kupari, P., Reinikainen, P., & Arrffman, I. (2002). *The Finnish Success in PISA - and some reasons behind it*. Jyvaskyla: Institute for Educational Reseach, University of Jyvaskyla.

Warm, T. A. (1985). Weighted maximum likelihood estimation of ability from Item Response Theory with tests of finite length. Oklahoma City: U.S. Coart Guard Institute.

Astron B38960 12-04