The Irish Journal of Education, 2006, *xxxvii,* pp. 101-119.

THE PISA ASSESSMENT OF SCIENTIFIC LITERACY

Eemer Eivers
Educational Research Centre
St Patrick's College, Dublin

Declan Kennedy
Education Department
University College Cork

Aspects of the performance of Irish 15-year olds on a test of scientific literacy in the OECD Programme for International Student Assessment (PISA) are described. The conceptualization of scientific literacy is discussed and related to how it was assessed in PISA 2003. The performance of Irish students, which was significantly above the OECD average, is described, and related to the performance of Irish students in the 2000 cycle of PISA. In describing students' performance, the lack of significant gender differences and the significant differences in relation to socioeconomic status, school sector, and student uptake of science are considered.

Results of previous international studies of science achievement have been mixed for Irish students. For example, in the 2000 cycle of PISA, Irish 15-year olds scored significantly above the average of the 31 countries that participated, and Ireland ranked 9th of the 27 participating OECD countries (Shiel, Cosgrove, Sofroniou, & Kelly, 2001). This is quite different from the results of the earlier Second International Assessment of Educational Progress (IAEP2), in which Irish students had the lowest mean for 9-year olds and the second lowest for 13-year olds (Martin, Hickey, & Murchan, 1992). The Third International Mathematics and Science Study (TIMSS) data were slightly more positive. At primary level, fourth class pupils achieved a mean science score that did not differ significantly from the overall OECD mean, ranking 10th of 17 OECD countries. At post-primary level, the mean of second year students ranked 9th of 17 OECD countries, and did not differ significantly from the OECD mean (OECD, 1997). A number of possible reasons for the variation in the performance of Irish students (as well as of students of other nationalities) in international comparative studies have been proposed. These include sampling on the basis of grade or age, variation in exclusions and participation rates, variation in data analysis and scoring procedures, and differences in the content areas assessed (O'Leary, Kellaghan, Madaus, & Beaton, 2000). It has also been suggested that country rankings in science achievement may be less stable than mathematics achievement, and more influenced by differences in the content areas that are assessed.

In this paper, following a description of how scientific literacy is defined in PISA, we examine the performance of Irish students in the 2003 PISA cycle to see if the positive results of the 2000 assessment still hold. Factors associated with performance on the literacy test will be identified.

DEFINING SCIENTIFIC LITERACY

PISA differs from previous international studies such as TIMSS[1], in that it sets out to assess how successful students have been in acquiring basic skills for adult life. It attempts not only to assess what students know, but also their ability to reflect upon what they know, and to apply their knowledge to real-life situations. Fifteen-year olds are the target of the assessment because at this age students are nearing the end of basic schooling in most countries and, following it, their educational experiences tend to diverge.

By directly testing for knowledge and skills close to the end of basic schooling, OECD/PISA assesses the degree of preparedness of young people for adult life and, to some extent, the effectiveness of education systems. Its ambition is to assess achievement in relation to the underlying objectives (as defined by society) of education systems, not in relation to the teaching and learning of a body of knowledge (OECD, 2003, p.14).

The notion of preparedness for adult life underpins each of the three domain frameworks that guide the design and implementation of PISA (reading, mathematics, and science literacies) and is apparent in its definition of scientific literacy:

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

In developing the framework for scientific literacy, a variety of descriptions and models of literacy and of science teaching were considered. However, given the emphasis on the application of knowledge in real-life situations, it is not surprising that the eventual framework (OECD, 2003) was heavily influenced by STS (Science-Technology-Society), or 'context-based', approaches. STS has a variety of interpretations. For example, Aikenhead (1994) has identified eight categories of STS teaching, ranging from using STS ideas to motivate pupils in traditional science courses, through courses that use STS content as a

[1] When Ireland participated in TIMSS in 1995, the acronym stood for the Third International Mathematics and Science Study. It has subsequently been renamed the Trends in International Mathematics and Science Study to allow more recent cycles to retain the acronym.

starting point for the development of scientific ideas, to courses that are wholly based on STS approaches. What STS materials and courses appear to have in common is the promotion of scientific literacy by developing students' understanding of one or more of the following:

- what is meant by science and technology, and how they relate to each other;
- the ways in which science and technology affect society;
- the way scientists work;
- the nature of science.

STS approaches also promote the discussion of personal opinions and values.

In Ireland, STS and related concepts have influenced science teaching at senior cycle in post-primary schools, as exemplified in the most recent Physics, Chemistry, and Biology syllabi which include topics designed to show students the links between the science they study in school and their everyday lives. Although the revised Junior Certificate science syllabus was not in place at the time of PISA 2003, the earlier syllabus also subscribed to many of the principles of STS, particularly the practical application of science to everyday life.

Arguments for the adoption of STS approaches to science curricula and teaching tend to fall into three main categories. The first, based on 'citizen science', argues that knowing about science helps citizens to think and act appropriately when dealing with scientific matters which may affect their lives and the lives of other members of the local, national, and global community. The second argument – 'relevant science' – claims that when science emphasizes applications rather than abstractions, interest levels are likely to be higher. Finally, the 'added-value' argument claims that approaches to science teaching that include decision-making and problem-solving may enhance students' more general skills in these areas [though this claim has been questioned (e.g., Millar, 1989)]. Given that the OECD is primarily an economic body, it is not surprising that constructs such as citizen science and added-value proved appealing. Both concur with the OECD's ideological emphasis on 'human capital' and on lifelong learning to facilitate optimum productivity.

FROM DEFINITION TO ASSESSMENT CONTENT

Once scientific literacy was defined, the next step was to clarify what lay within the domain, thereby helping to frame the contents and methodology of the assessment. The definition itself was built on three distinct dimensions – s*cientific knowledge/concepts*, *scientific processes*, and the *contexts* in which scientific knowledge or scientific processes are assessed. Each of these was considered in the construction of test items, and items were classified according to each one.

Scientific Knowledge/Concepts

As science was not a major domain in 2003, the development of assessment items was somewhat constrained by the need to produce a relatively small number of items that adequately represented the main branches of science, described as physics, chemistry, biology, and earth and space science. Other criteria for the inclusion of items included the relevance of the concept to daily life and its projected relevance over the next 10 years, and whether the concept could be combined with selected scientific processes. Consideration of these factors led to the identification of 13 major scientific 'themes', which are listed in Table 1.

Table 1
The 13 Major 'Themes' Identified in the PISA Framework for Scientific Literacy

Structure and properties of matter (thermal and electrical conductivity)
Atmospheric change (radiation, transmission, pressure)
Chemical and physical changes (state of matter, rates of reaction, decomposition)
Energy transformations (energy conservation, energy degradation, photosynthesis)
Forces and movement (balanced/unbalanced forces, velocity, acceleration, momentum)
Form and function (cell, skeleton, adaptation)
Human biology (health, hygiene, nutrition)
Physiological change (hormones, electrolysis, neutrons)
Bio-diversity (species, gene-pool, evolution)
Genetic control (dominance, inheritance)
Ecosystems (food chains, sustainability)
The earth and its place in the universe (solar system, diurnal and seasonal changes)
Geological change (continental drift, weathering)

Scientific Processes

Processes are the actions that are used to obtain, interpret, and use evidence to gain knowledge. '… [T]hey become scientific processes when the subject matter is drawn from scientific aspects of the world and the outcome of using them is to further scientific understanding' (OECD, 2003, p. 136). The PISA 2003 framework outlines three processes within the science domain:

- describing, explaining, and predicting scientific phenomena;
- understanding scientific investigation;
- interpreting scientific evidence and conclusions.

Context

Given that one of the underlying themes of PISA is that it should assess how students can *use* science, rather than assess how much of a science curriculum they have learned, an important element of the assessment is the context in which scientific knowledge and processes are applied. Three broad context categories were defined:

- science in life and health;
- science in earth and environment;
- science in technology.

The first category (*science in life and health*) includes issues such as health and nutrition, and the interdependence of biological systems. *Science in earth and environment* includes issues such as pollution, soil production, and climate, while *science in technology* includes biotechnology, energy use, and transportation.

Assessment Items

Each PISA test booklet is composed of *units* of material, which are groups of items (questions) about a common topic. The final PISA 2003 assessment contained 13 science units, with 35 questions. Of the final item pool, 25 items from 10 units had also been used in the 2000 assessment. A variety of formats was used, such as simple multiple-choice (students circle one of four options), closed constructed response (students write a short, simple answer that is compared to a single correct answer), and open constructed response items (where students write a more detailed response that requires complex marking by trained markers). Most of the questions were either simple multiple-choice (37%) or open constructed response (40%). Twenty percent were complex multiple-choice, while 3% were short-response items.

Of the 13 themes identified in the framework, the theme most commonly represented in questions was *the earth and its place in the universe* (20% of items were related to this theme), followed by *the structure and property of matter* (17%), *energy transformations* (11%), and *physiological change* (11%). When items were classified by process, almost half related to *describing, explaining, and predicting scientific phenomena*; 31% to *understanding scientific investigation;* and 20% to *interpreting scientific evidence and conclusions.* Finally, each of the three contexts was represented in approximately one-third of questions. Information on the precise distribution of items is detailed in the PISA framework (OECD, 2003). Due to the small item set, only an overall score was available for science in 2003. However, as science will be the main focus of the 2006 assessment, subscales and proficiency levels will then be available.

SAMPLE PISA 2003 SCIENCE ITEMS

The following examples are from the PISA 2003 assessment from a unit called Cloning. This unit was also included in the 2000 assessment, and was released for review after the 2003 assessment. Students were asked to read a medium-length passage (Figure 1) about Dolly, the first cloned sheep, and then to answer three questions about the passage. In the test booklets, the unit was accompanied by a large photograph of Dolly (not included here). The items in the unit deal with the process describing, explaining, and predicting scientific phenomena, in the context of science in life and health. Each of the items shown can be placed on an item difficulty scale, which is the same as the student performance scale (mean of 499.6 and standard deviation of 105.5). Thus, for example, an item with a score of 550 is of medium difficulty, while a score of 650 represents a very difficult item.

Figure 1
Sample Stimulus and Questions from the PISA 2003 Assessment of Scientific Literacy

Cloning
Read the newspaper article and answer the questions that follow.
A copying machine for living beings?

Without any doubt, if there had been elections for the animal of the year 1997, Dolly would have been the winner! Dolly is a Scottish sheep that you see in the photograph. But Dolly is not just a simple sheep. She is a clone of another sheep. A clone means ‘a copy’. Cloning means copying ‘from a single master copy’. Scientists succeeded in creating a sheep (Dolly) that is identical to a sheep that functioned as a ‘master copy’.

It was the Scottish scientist Ian Wilmut who designed the ‘copying machine’ for sheep. He took a very small piece from the udder of an adult sheep (sheep 1).

From that small piece he removed the nucleus; then he transferred the nucleus into the egg-cell of another (female) sheep (sheep 2). But first he removed from that egg-cell all the material that would have determined sheep 2 characteristics in a lamb produced from that egg-cell. Ian Wilmut implanted the manipulated egg-cell of sheep 2 into yet another (female) sheep (sheep 3). Sheep 3 became pregnant and had a lamb: Dolly.

Some scientists think that within a few years it will be possible to clone people as well. But many governments have already decided to forbid cloning of people by law.

Question 1. Which sheep is Dolly identical to?

Response Option	Percent choosing each response	
	Ireland	OECD Average
Sheep 1*	58.7	64.7
Sheep 2	19.8	13.5
Sheep 3	16.1	15.4
Dolly's father	5.0	4.9
Missing	0.5	1.5

*=Key

Question 2. In line 14 the part of the udder that was used is described as 'a very small piece'. From the article text you can work out what is meant by 'a very small piece'. That 'very small piece' is

Response Option	Percent choosing each response	
	Ireland	OECD Average
a cell*	48.4	48.7
a gene	13.2	17.5
a cell nucleus	27.5	25.1
a chromosome	10.4	7.4
Missing	0.5	1.3

*=Key

Question 3. In the last sentence of the article it is stated that many governments have already decided to forbid cloning of people by law. Two possible reasons for this decision are mentioned below. Are these reasons scientific reasons? Circle either 'Yes' or 'No' for each.

Reason:	Scientific?
Cloned people could be more sensitive to certain diseases than normal people.	Yes* / No
People should not take over the role of a 'Creator'.	Yes / No*

*=Key

Percent answering ...	Ireland	OECD Average
Both correctly	71.3	62.1
One correctly	13.8	18.8
None correctly	14.0	18.2
Missing	0.8	1.0

The first sample question is a multiple-choice item that deals with the theme of genetic control. Students are asked to which sheep Dolly is identical, and are provided with four possible response options, one of which they must circle. The item is relatively straightforward; 59% of students in Ireland and 65% of students overall answered correctly. Placed on an item difficulty scale, this item was assigned a score of 494.

The second sample question is another multiple-choice item, dealing with the theme of form and function. It proved to be slightly more difficult, as students must show an understanding of cell structure. On the item difficulty scale, question 2 was assigned a score of 572. Just under half of students (in Ireland and overall) answered the item correctly.

The third and final example is a complex multiple-choice item that requires students to show that they can differentiate between statements that are and are not scientifically based. Students were presented with two reasons – one scientific, the other religious – why governments might oppose cloning, and were asked to indicate if each was scientifically based. Those who correctly labelled both statements were given full credit. The item difficulty level was 507, meaning that it was of medium difficulty. At 71%, the percentage of students in Ireland who answered both parts of the item correctly was 9% higher than the OECD average of 62 percent. Girls were more likely than boys to answer correctly. For example, in Ireland, 75.9% of girls answered correctly, compared to 66.8% of boys. This may be compared with the OECD average scores for boys and girls on questions 1 and 2 which were no more than a percentage point apart. The largest Irish gender difference on these items was on question 2 ('very small piece'), which 50.2% of girls and 46.6% of boys answered correctly.

It is notable that the level of missingness for these three examples is extremely low, meaning that almost all students attempted them. This does not hold true for a significant number of PISA items. The sample items presented are multiple-choice, and for items in this format, missingness tended to be low. For open-constructed responses (often requiring a quite detailed answer from the student), missingness tended to be considerably higher. Readers who wish to review some sample constructed response items are directed to *http://www.erc.ie/pisa/* or to Cosgrove, Shiel, and Kennedy (2002).

PERFORMANCE ON THE 2003 ASSESSMENT

The Overall Performance of Irish Students on Scientific Literacy

Irish students achieved a mean score of 505.4 on the scientific literacy scale (OECD mean = 499.6; standard deviation = 105.5) (Table 2). The Irish mean is the 16th highest of the 40 participating countries, and the 13th highest of the 29 OECD countries for which reliable achievement data were available[2]. Students in both Finland and Japan obtained mean scores that were over 40 points (more than two-fifths of a standard deviation) higher than students in Ireland, while students in three countries (Tunisia, Brazil, Indonesia) obtained mean scores that were at least one full standard deviation lower than the Irish mean.

Table 2

Mean Achievement Scores and Standard Deviations on the PISA 2003 Scientific Literacy Scale, by Country

	Country	Mean	SD		Country	Mean	SD
Science score significantly **higher** than Ireland's	Finland	548.2	90.8	Science score significantly **lower** than Ireland's	Iceland	494.7	95.6
	Japan	547.6	109.4		U.S.	491.3	101.6
	Hong-Kong-Ch	539.5	94.1		Austria	491.0	97.0
	Korea	538.4	100.5		*Russian Fed*	489.3	99.8
	Liechtenstein	525.2	103.5		*Latvia*	489.1	92.7
	Australia	525.1	101.8		Spain	487.1	100.2
	Macao-Ch	524.7	87.9		Italy	486.5	107.8
	Netherlands	524.4	98.5		Norway	484.2	103.8
	Czech Rep.	523.3	100.6		Luxembourg	482.8	102.8
	New Zealand	520.9	104.0		Greece	481.0	100.6
	Canada	518.7	99.1		Denmark	475.2	101.8
Science score not significantly different to Ireland's	Switzerland	513.0	107.5		Portugal	467.7	93.4
	France	511.2	110.8		*Uruguay*	438.4	109.1
	Belgium	508.8	107.4		*Serbia & M.*	436.4	82.7
	Sweden	506.1	106.8		*Turkey*	434.2	95.9
	IRELAND	**505.4**	**93.0**		*Thailand*	429.1	81.3
	Hungary	503.3	97.3		Mexico	404.9	86.7
	Germany	502.3	111.4		*Indonesia*	395.0	68.0
	Poland	497.8	102.4		*Brazil*	389.6	98.3
	Slovak Rep	494.9	102.2		*Tunisia*	384.7	87.3

Shaded countries are those whose means are significantly higher than the OECD mean (Mean 499.6; SD: 105.5).
Source: Cosgrove, Shiel, Sofroniou, Zastrutski, & Shortt (2005), Table 3.24.

[2] As the UK did not meet the required sampling response rate standards, its data are not included in the table.

A simple ranking of each country by mean score does not take into account associated standard errors (an estimate of the extent to which a country's obtained mean may be expected to vary around the 'true' mean for that country). When this is done using Bonferroni multiple comparisons, 11 countries were found to have significantly higher mean scientific literacy scores than Ireland, 20 have significantly lower scores, and a further eight do not differ significantly from Ireland (see OECD, 2004, Figure 6.10).

Where large numbers of comparisons are made – as in PISA, where the mean scores of a large number of countries are compared – Bonferroni adjustments can lead to very conservative interpretations of significance. Therefore, a technique called nonparametric maximum likelihood (NPML) estimation was also used to group countries' performance [see Cosgrove et al. (2005) for more details]. Using NPML, six groups of countries are distinguishable, with four countries (Finland, Japan, Hong Kong-China, and Korea) in the top-scoring group. Ireland is in the third group, meaning that the Irish score is very similar to scores obtained in Belgium, Sweden, Hungary, and Germany.

Comparing Performances in 2000 and 2003

In PISA 2000, the mean scientific literacy score obtained by Irish students was 513.4 (on a scale which had a mean of 500 and a standard deviation of 100). Although Ireland's score of 505.4 in the 2003 assessment is lower than the mean score obtained in 2000, the difference is not statistically significant. In terms of country rankings, Ireland obtained the 9th highest score in 2000, a score that is significantly lower than that obtained in six countries and significantly higher than that obtained in 17 countries.

Table 3 lists all the countries that took part in both the 2000 and 2003 assessments, grouped by whether their mean scientific literacy scores are significantly higher, similar to, or significantly lower than Ireland's in 2003. Countries whose grouping had changed since 2000 are highlighted in bold font, and an accompanying arrow shows the direction in which they had moved. For example, of the 14 countries with significantly lower mean science scores than Ireland in 2003, three (Austria, Norway, USA) are highlighted and have an arrow pointing down. This means they had moved from not differing significantly from Ireland in 2000 to the lower group in 2003.

Table 3
Countries Grouped by Whether or Not their Mean Scores Differed Significantly from Ireland's in 2003, with Changes (if any) from 2000 Grouping Indicated

Higher	No significant difference	Lower
Australia ⇑	**Belgium** ⇑	⇓ **Austria**
Canada	France	*Brazil*
Czech Rep ⇑	**Germany** ⇑	Denmark
Finland	**Hungary** ⇑	Greece
Japan	Ireland	Iceland
Korea	**Poland** ⇑	Italy
Liechtenstein ⇑⇑	Sweden	*Latvia*
New Zealand	**Switzerland** ⇑	Luxembourg
		Mexico
		⇓ **Norway**
		Portugal
		Russian Federation
		Spain
		⇓ **USA**

Bold denotes a country whose grouping relative to Ireland had changed since 2000; arrow denotes direction of change.

In the case of countries that did not differ significantly from Ireland in 2003, only France and Sweden also belonged to this group in 2000. Belgium, Germany, Hungary, Poland, and Sweden all obtained significantly lower scores than Ireland in 2000, but in 2003 their score improved sufficiently for them to move into the middle grouping. Of the eight countries shown with significantly higher mean scores than Ireland in 2003, three (Australia, Czech Republic, Liechtenstein) did not fall into this group in 2000. Indeed, Liechtenstein moved from the group with significantly lower mean scores than Ireland in 2000 to the group with significantly higher mean scores in 2003.

Apart from Liechtenstein, the most notable gains were made by Luxembourg[3], the Russian Federation, and Latvia, where mean scores increased

[3] There was a major difference in test conditions in Luxembourg in the 2000 and 2003 assessments. In 2000, students were not allowed to choose their test language; in 2003, they were.

by at least 30 points since 2000. In contrast, the mean score in Austria was almost 28 points lower in 2003 than in 2000. Amongst these changes, some constants remain. Finland, Japan, and Korea performed well above average in both years; Ireland performed just above the OECD average; and, Brazil, Mexico, and Portugal performed well below average.

Performance of Low and High Achievers

One method of describing the performance of high and low achievers is in terms of performance at 'benchmark' percentiles. For the purpose of this paper, students scoring at the 10th and 90th percentiles were taken as representative of low- and high-achieving students, respectively. For Ireland in 2003, the difference in science scores between students scoring at these benchmarks is 240.6 points, which is smaller than the OECD average difference of 272.6 points, and means that the distance between low- and high-achieving students in Ireland is less than in most participating countries. Ireland's score of 383.9 at the 10th percentile is the eleventh highest at that benchmark, and is considerably higher than the OECD average of 361.6. However, Ireland's score of 624.5 at the 90th percentile is the 20th highest, and is lower than the OECD average (635.7). Indeed, Ireland's scores are higher than the OECD average at the 5th, 10th, and 25th percentiles, but lower at the 75th, 90th, and 95th percentiles. This means that while Ireland's low-achieving students do well, compared to other countries, its high-achieving students do not.

When compared to performance in 2000, some similarities emerge. In both years, the difference in science scores between students scoring at the 10th and 90th percentiles was lower in Ireland than the OECD average, and Irish students scoring at the 10th and 25th percentiles obtained scores above the OECD average. However, whereas in 2000 Ireland's high-performing students (i.e., those at the 90th percentile) scored above the OECD average and were ranked 10th highest, in 2003 Ireland's high performers were below the OECD average. This is not because the performance of Irish students at the 90th percentile deteriorated, but because of a significant improvement between 2000 and 2003 in overall performance in other countries at the top end of the science scale.

FACTORS ASSOCIATED WITH SCIENTIFIC LITERACY SCORES

In this section, performance on the science achievement test is linked to contextual data, largely derived from School and Student Questionnaires. More detail on each of the points discussed (and additional statistical details such as standard errors and confidence intervals) is available in Cosgrove et al. (2005).

Studying Science in Junior Cycle

In Ireland, 9.9% of students assessed in PISA 2003 had not studied junior cycle science. Female uptake was lower than male uptake: 14.6% of girls, but only 5.2% of boys, did not study science. The rate of science uptake increased only marginally since 2000, when 11.2% did not take the subject. Students who did not study science obtained a mean score of 451.8. This is significantly lower than the mean score of 547.1 obtained by students who studied Higher level science (a difference of 95 points), and higher, but not significantly higher, than the mean score of 433.3 obtained by students who studied Ordinary level science. A similar pattern was observed in 2000.

Student Socioeconomic Status (SES)

Maternal and paternal occupations were classified using the International Socioeconomic Index (ISEI) scale developed by Ganzeboom, de Graaf and Treiman (1992). Although ISEI is a continuous scale, for reporting purposes students were divided into those whose parents were categorized as of either low, medium, or high SES. SES was significantly related to scientific literacy scores, with students in the high SES group (mean score of 542.5) outperforming students in the medium group (509.6) and in the low group (470.8). Similar differences were found in the 2000 assessment.

Gender Differences

Irish data for 2000 and 2003 reveal no significant gender differences for mean scores on the scientific literacy scale or at key percentile points (e.g., 10th and 90th percentiles). By comparison, the OECD average for 2003 showed a slightly, but significantly, higher score for girls than for boys (a 5.8 point difference), but with no clear pattern of gender difference across countries. Thus, science appears to contrast with the more ‘gendered’ domains of reading (girls outperformed boys in all but one country in 2003) and mathematics (boys outperformed girls in 21 of 29 OECD countries in 2003).

School Characteristics

The mean score obtained by students in schools designated as disadvantaged (478.6) is significantly lower than the mean score of students in non-designated schools (515.2). Although the difference of 36.3 points between the two groups is lower than the 48.2 point difference in 2000, the reduction does not reach statistical significance. Schools were also categorized into those with a low, medium, or high weighted percentage of students entitled to a Junior Certificate examination fee waiver. Students enrolled in schools where a relatively small percentage of students were entitled to a waiver obtained a significantly higher mean scientific literacy score (537.3) than students in 'medium fee waiver' (508.3) and 'high fee waiver' schools (470.6). There was also some variation in mean score by school sector. The mean score of 518.7 achieved by students enrolled in secondary schools is significantly higher than the score of 498.7 of students in community/comprehensive schools, and the score of 473.4 of students in vocational schools. A similar pattern was found in 2000.

Analysis of student performance with reference to school gender composition revealed that students in all boys' schools obtained a significantly higher mean score (528.0) than students in all girls' schools (516.2), in mixed but primarily boys' schools (500.4), and in mixed but primarily girls' schools (479.8). This is slightly different from the performance of students in PISA 2000 when students in all boys' schools obtained a significantly higher mean score than students in mixed schools. The gap between students in all girls' and all boys' schools was 6.2 points, smaller than the 11.8 point gap in 2003.

As well as disadvantaged status, schools were categorized by an index of Economic, Social and Cultural Status (ESCS) which was based on student responses (aggregated to the school level) to items about parental educational attainment, occupation, and number of home possessions, including books and 'cultural' objects. Students in high ESCS schools obtained a significantly higher mean science score (537.9) than students in medium (509.6) and low (467.0) ESCS schools. Indeed, the gap in mean scores (70.9 points) between students in low and high ESCS schools is much larger than the gap between students in schools designated or not designated as disadvantaged, or in different school sectors.

EXPLAINING ACHIEVEMENT IN SCIENCE

Many of the school and student characteristics considered in the preceding section are interrelated, making effects difficult to disentangle. For example, a school's designated disadvantaged status, the percentage of its students who are entitled to a fee waiver, and the school-level ESCS score are interrelated. For this reason, a hierarchical linear model, which permits simultaneous examination of the effects of variables and can distinguish between the effects of school- and student-level variables, was developed (Cosgrove et al., 2005).

Two school-level variables survived in the final model: a measure of the school's disciplinary climate and the percentage of students in a school who were entitled to the Junior Certificate examination fee waiver. The student-level variables that survived were gender, socioeconomic status, lone-parent family status, number of siblings, measures of the number of books in the student's home and of home educational resources, the student's level of absenteeism, current grade level, and whether or not the student had studied science.

As an example of how the variables were linked to achievement after controlling for other variables, students who had studied science tended to score 38 points higher on scientific literacy than students who had not. Students with one sibling tended to obtain higher scores than students with none or more than one sibling, while students in the high SES category would be expected to score approximately 33 points higher than students in the low SES category. Effects were also marked for the number of books in students' homes: students with no books at home tended to score approximately 71 points lower than students with more than 500 books. Overall, the variables included in the model accounted for 80.2% of between-school variance and 31.2% of within-school variance.

CONCLUSION

A number of factors might lead to an expectation of average performance by Irish students on a test of scientific literacy. For example, concerns have been raised repeatedly about the uptake of science by post-primary school students, particularly at senior cycle, and about the fact that not all junior cycle students take science as a subject (e.g., Task Force on the Physical Sciences, 2002). Furthermore, earlier international studies that assessed science knowledge reported mixed results for Irish students (e.g., Martin et al., 1992; OECD, 1997; Shiel et al., 2001). Thus, the fact that Irish students performed slightly above the OECD average on scientific literacy in PISA 2003, maintaining the above average performance of the 2000 assessment, must be regarded as welcome. However, some aspects of the results merit further examination.

While Ireland has performed better on the PISA assessments than on previous international assessments of science achievement, a question that arises is whether the improvement reflects a genuine raising of standards of science knowledge, or is an artefact of assessment methodology, design, or test content. One possible explanation is that PISA, in contrast to earlier studies, attempted to measure 'scientific literacy' rather than how much science students had learned in school.

However, the extent to which the PISA science (and, indeed mathematics) tests measure scientific (and mathematical) literacy, as opposed to 'the ability to apply mathematical and scientific knowledge in literary contexts' is debatable (Smithers, 2004, p. 7). Certainly, there is a very heavy reading load on some of the science items, leading Shiel et al. (2001) to argue that language may have been a construct-irrelevant factor contributing to performance. Given that Irish students performed well on the reading literacy test, it is possible that their scientific literacy scores were boosted by their reading ability.

Further, while PISA is intended to be independent of curricula, the extent to which this can be achieved is also debatable, given that any assessment will always match curricula in some countries more closely than in others. Rocher's (2003) analyses of PISA 2000 data suggest that the curricula of countries are an important source of bias in how countries perform. In the case of Ireland, a study that linked PISA 2000 science items[4] to the Junior Certificate science curriculum in use at that time found that over half of items assessed topics covered in the curriculum, while over 90% assessed processes with which students would be expected to be familiar (Shiel et al., 2001). While the context of approximately 80% of items was unfamiliar, this was largely because reading through large units of text, extracting relevant and discarding irrelevant information, was not a major element of the science curriculum. Thus, apart from the heavy reading load and analytic element, it could be argued that the curriculum followed by Irish students meant they were reasonably familiar with many aspects of the PISA science assessment.

The finding that Irish boys and girls did not differ significantly in their mean scientific literacy scores is not consistent with Junior Certificate science examination results. In the 2003 examination (when a majority of students who participated in the PISA assessment sat the Junior Certificate), girls scored approximately half a grade point higher than boys (9.43 versus 8.95, respectively)[5]. Further, there were no gender differences on PISA scientific literacy at key percentile points, even though, for example, 48% of girls, but only 40% of boys, taking Higher level science obtained an A or B grade.

The sizeable minority of girls who did not take science in junior cycle might be perceived to be responsible for Irish girls' poorer performance on PISA than on the Junior Certificate examination. However, this ignores the findings of the multilevel model of science achievement that, when other factors were taken into account, girls' scientific literacy scores still tended to be lower than those of boys. Further research is needed to determine if this reversal of gender

[4] Similar analyses are not available for the 2003 science items. However, most of the 2003 science items had previously been used in 2000, suggesting that many of the points relating to the 2000 item set also apply to the 2003 item set.

[5] Junior Certificate examination results were mapped on to a 9-point scale ranging from scores of 4 (for F in Ordinary level science) to 12 (A in Higher level science). The Ordinary and Higher levels overlapped at 9 (equivalent to A on Ordinary and D on Higher level).

differences reflected in public examinations is due to the structure or format of the assessment[6], to factors such as the use of simulated scores for students who did not take the science element of the assessment, or to attitudinal differences towards the assessment.

The finding that 10% of students (almost 15% of girls) did not study science at junior cycle level needs careful consideration. Despite recent efforts to increase awareness of science and uptake [for example, programmes such as STEPS (Science, Technology and Engineering Programme for Schools) and The Science Bus], there has been minimal increase in uptake since the 2000 assessment. The data in 2000 and 2003 indicate that students who studied Higher level science obtained a higher mean scientific literacy score than students who did not. This is hardly surprising, and serves to underline the importance of studying science in order to be able to gain and use scientific knowledge. However, the mean scores of students who studied Ordinary level science and of students who did not are not significantly different. This suggests that precisely what Ordinary level science students learn needs closer inspection.

The next PISA assessment, which is scheduled to take place in 2006, will differ from its predecessors in two major aspects. First, most Irish students will have been taught using the revised Junior Certificate science syllabus (introduced in most Irish schools in September 2003). The revised syllabus is heavily influenced by the STS movement, and is intended to have a much stronger focus on the interpretation and use of science in real life situations than was evident in the older syllabus. It will be important to monitor if greater symmetry between syllabus and PISA science leads to any change in the performance of Irish students on PISA. Secondly, since science will be the main domain, or major focus, of the assessment in 2006, the greatly increased number of items should allow examination of the effects of various item characteristics, such as the volume of reading required for the stimulus text and question, the influence of the scientific processes covered, and the effects of different types of context. It should also be possible to separate the effects of scientific knowledge from the ability to apply scientific processes, to describe student achievements along proficiency levels similar to those that already exist for mathematics and reading, and to compare student performance along a number of distinct science themes and content areas.

[6] For example, Bolger & Kellaghan (1990) found that boys do better than girls on multiple-choice items, while girls do better than boys on ‘free-response’ items.

REFERENCES

Aikenhead, G. (1994). What is STS teaching? In J. Solomon & G. Aikenhead (Eds), *STS education: International perspectives on reform*. New York: Teachers College Press.

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

Cosgrove, J., Shiel, G., & Kennedy, D. (2002). The performance of Irish students in scientific literacy in the Programme for International Student Assessment (PISA). *Irish Journal of Education, 33,* 53-70.

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

Ganzeboom, H.B.G., deGraaf, P.M., & Trieman, D.J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21,* 1-56.

Martin, M.O., Hickey, B.L., & Murchan, D.P. (1992). The Second International Assessment of Educational Progress: Mathematics and science findings in Ireland. *Irish Journal of Education, 26,* 5-146.

Millar, R. (1989). What is 'scientific method' and can it be taught? In J. Wellington (Ed.), *Skills and processes in science education* (pp. 47-62). London: Routledge.

OECD (Organisation for Economic Co-operation and Development). (1997). *Education at a glance: OECD indicators.* Paris: Author.

OECD. (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills.* Paris: Author.

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

O'Leary, M., Kellaghan, T., Madaus, G., & Beaton, A. (2000). Consistency of findings across international surveys of mathematics and science achievement: A comparison of IAEP2 and TIMSS. *Education Policy Analysis Archives, 8,* no. 43. *http://epaa.asu.edu/epaa/v8n43.html*

Rocher, T. (2003). *On the methods used for international assessments of educational competences.* Paris: Office of Statistical Studies, French National Ministry of Education. (Translated by Jason Tarsh, Analytical Services, UK Department for Education and Skills, London).

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

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

Task Force on the Physical Sciences. (2002). *Report and recommendations. http://www.skillsireland.ie/press/relatedreports/index.html*