

Does Student Engagement Explain Performance on PISA?

Comparisons of Response Patterns on the PISA Tests Over Time

Jude Cosgrove

Educational Research Centre, St Patrick's College, Dublin 9

December 2011

Does Student Engagement Explain Performance on PISA?

Comparisons of Response Patterns on the PISA Tests Over Time

Jude Cosgrove

Table of Contents

Introduction	3
Aims of the Present Study.....	4
Method	4
<i>Selection of Items for Analysis</i>	4
<i>Analysis Strategy</i>	8
Results.....	10
<i>Trends in Achievement in Ireland, OECD and Comparison Countries</i>	10
<i>Analyses of Reading Items, 2003 and 2009</i>	11
<i>Trends in the Response Patterns of Sub-Groups in Reading, 2003 and 2009</i>	22
<i>Are the Response Patterns for Link Reading Items the same as for New Reading Items?</i>	27
<i>Analyses of Mathematics Items, 2003 and 2009</i>	27
<i>Analyses of Science Items, 2006 and 2009</i>	35
Summary and Conclusions.....	39
References	44

Introduction

PISA, or the Programme for International Student Assessment, is an initiative of OECD member states. It is a three-yearly survey of the knowledge and skills of 15-year-old students. One of the objectives of PISA is to provide information on trends in achievement over time. In PISA 2009 in Ireland, the results for reading literacy and, to a lesser extent, mathematical literacy, indicated a marked decline in performance relative to previous PISA surveys dating back to 2000. These results attracted media attention and commentary, and were presented to the Joint Oireachtas Committee on Education and Skills in January 2011. They were also the subject of a debate of that Committee in May 2011¹.

In response to these results, the Department of Education and Skills requested an independent review of the PISA 2009 results by Statistics Canada, and the Educational Research Centre also conducted a detailed internal review. Results of these reviews are in Cosgrove et al. (2010), Shiel et al. (2010), and LaRoche and Cartwright (2010), and summarised in the PISA 2009 summary report for Ireland (Perkins et al., 2010) and in the Teachers' Guide to the PISA 2009 results (Perkins et al., 2011).

The reviews that were conducted highlighted a need to further analyse students' responses to the PISA assessments. It had been noted (Cosgrove et al., 2010, pp. 28-29; LaRoche & Cartwright, 2010, pp. 4-5; p. 32) that students in Ireland appeared to be disengaged from the PISA 2009 print assessments to a greater degree than in previous cycles. This was evidenced in their behaviour during some of the testing sessions observed and also in the percentages of test questions that they did not attempt. Further, it was not possible to establish, at the time of writing of these reports, whether students' levels of engagement were the same on the digital reading assessment as they were on the print assessment; however, it was thought that this may not have been the case since students in Ireland had a mean score that was some 13 points higher on the digital reading assessment than on the print reading assessment (e.g. Cosgrove et al., 2011).

The present report is one of three which provides a more in-depth look at students' response patterns on the PISA tests. It examines student response patterns over successive PISA cycles, focusing on reading and mathematics. The second report (Cosgrove & Moran, 2011) examines students' response patterns on the print and digital assessments in PISA 2009. Cartwright (2011) has also examined students' response patterns in Ireland across PISA cycles with reference to international patterns. Together, the present report along with information documented in Cosgrove and Moran (2011) and Cartwright (2011) give a comprehensive picture of students' response patterns on the PISA assessments. The three reports are summarised in Chapter 9 of the PISA 2009 national report (Perkins et al., 2012).

It is assumed that readers are familiar with the aims, design and main results of PISA. Readers who wish to familiarise themselves with the aims and design of PISA as well as the main results are referred to OECD (2010a, b, c, d, e; 2011), as well as to Perkins et al. (2011). Details of PISA publications can also be found at www.erc.ie/pisa and www.pisa.oecd.org.

¹ <http://debates.oireachtas.ie/EDJ/2011/01/13/00004.asp>

Aims of the Present Study

In previous analyses of the reported decline in reading achievement in Ireland between 2000 and 2009, a decrease in percent correct and corresponding increase in percent of missing items was found. It was argued that Ireland's very large decline since 2000 (31 points, the largest of all countries for which comparisons can be made) could be due in part to a decrease in proficiency, engagement with the test, or some mixture of these. The same could be argued for mathematics since 2003 (which declined by 16 points) but not so for science (no change since 2006).

Aside from these factors, issues in the scaling of PISA, the methods that PISA uses to produce trends, demographic changes in the school-going population, and the way in which PISA was implemented in schools are considered elsewhere (LaRoche & Cartwright, 2010; Cartwright, 2011).

The analyses described in this document examine two possibilities:

1. The decline in PISA reading is due to a decrease in proficiency
2. The decline in PISA reading is due to a decrease in engagement.

Of course, some mixture of these could be possible. However, given that there is no corroborating evidence for a decline of the magnitude observed in Ireland between 2000 and 2009, particularly in the case of reading, and that declines in educational standards of this extent have never been recorded in this timeframe before, the hypothesis that the decline in reported PISA scores is due to a decrease in engagement in PISA more so than a decline in student proficiency is put forward in this paper. The analyses include all three assessment domains to examine whether differential changes in average achievement are due to differing response patterns across cycles.

Method

Selection of Items for Analysis

It was necessary to identify a common set of items across cycles that were administered in a manner (sequence) similar enough to allow comparisons of item responses.

The PISA test design is such that each student attempts a booklet consisting of four half-hour blocks, and since 2003, the test design has been balanced, implying that each block appears in each of the four positions. This is done to eliminate the confounding effect of test fatigue in the estimation of item difficulties since booklet data are pooled together. In PISA 2000, however, the test design was not balanced, i.e. not all blocks appeared in all positions. This makes comparisons of booklet position effects between 2000 and all other cycles inherently problematic. Furthermore, in selecting reading items to estimate trends, intact blocks were not selected from 2000 for use in 2003. However, since 2003, the same two intact blocks have been used. Both blocks are analysed here.

In the case of mathematics, just one common block was selected for analysis, and, since intact blocks were not selected from 2003, comparisons are limited to 2006 – 2009. In the case of science, intact blocks were not selected from 2006 to form the blocks used in 2009, so the analysis is more limited in that it involves comparing the same block within a cycle in positions 1 and 4, but blocks are not common *across* cycles.

During PISA test scaling, not reached items are scored as missing in the computation of item parameters, but are scored as zero (as are missing responses) in the computation of student scores. Therefore, increases in missing responses are likely to result in increases in item difficulty while increases in both missing and not reached responses are likely to result in a decline in student scores.

Table 1: Summary of item characteristics in reading block R2, 2003 and 2009

Order in unit	Item	Question Format	Process	Difficulty 1	Difficulty 2
1	R227Q01	Multiple Choice	Integrate and interpret	535	
2	R227Q02	Complex Multiple Choice	Access and retrieve	422	611
3	R227Q03	Open Constructed Response	Reflect and evaluate	540	
4	R227Q06	Short Response	Access and retrieve	459	
5	R111Q01	Multiple Choice	Integrate and interpret	488	
6	R111Q02B	Open Constructed Response	Reflect and evaluate	556	697
7	R111Q06B	Open Constructed Response	Reflect and evaluate	557	593
8	R055Q01	Multiple Choice	Integrate and interpret	401	
9	R055Q02	Open Constructed Response	Reflect and evaluate	554	
10	R055Q03	Open Constructed Response	Integrate and interpret	525	
11	R055Q05	Open Constructed Response	Integrate and interpret	448	
12	R104Q01	Closed Constructed Response	Access and retrieve	385	
13	R104Q02	Closed Constructed Response	Access and retrieve	619	
14	R104Q05	Short Response	Access and retrieve	581	774

Difficulty 1 and difficulty 2 refer to the PISA score point difficulties of each item. Partial credit items have two difficulty estimates.

Table 2: Summary of item characteristics in reading block R1, 2003 and 2009

Order in unit	Item	Question Format	Process	Difficulty 1	Difficulty 2
Deleted	R219Q01T	Open constructed response	Reflect and evaluate	N/A	
Deleted	R219Q01E	Open constructed response	Reflect and evaluate	N/A	
1	R219Q02	Open constructed response	Reflect and evaluate	413	
2	R067Q01	Multiple choice	Integrate and interpret	336	
3	R067Q04	Open constructed response	Reflect and evaluate	466	581
4	R067Q05	Open constructed response	Reflect and evaluate	466	511
5	R102Q04A	Open constructed response	Integrate and interpret	626	
6	R102Q05	Closed constructed response	Integrate and interpret	573	
7	R102Q07	Multiple choice	Integrate and interpret	389	
8	R220Q01	Short response	Access and retrieve	572	
9	R220Q02B	Multiple choice	Integrate and interpret	492	
10	R220Q04	Multiple choice	Integrate and interpret	495	
11	R220Q05	Multiple choice	Integrate and interpret	392	
12	R220Q06	Multiple choice	Integrate and interpret	476	

Difficulty 1 and difficulty 2 refer to the PISA score point difficulties of each item. Partial credit items have two difficulty estimates.

The content of the reading items used in the present study is shown in Tables 1 and 2, i.e. item IDs, question format, reading process assessed, and item difficulty expressed on the PISA scale with a mean of 500 and standard deviation of 100. In four instances, the items are partial credit, which implies that there are two item difficulties associated with these – one for a partially correct response, and another for a fully correct response. Table 3 shows the content of block M2 (as labelled in 2009)/M1 (2006), while Table 4 shows the characteristics of items for S4 (2006) and Table 5 shows these for block S1 (2009). (Items were deleted internationally from blocks S1 and S2, and S4 is in position 4 following three science clusters, so S4 was selected for these reasons.)

Table 3: Summary of item characteristics in mathematics block M2 (2009) / M1 (2006)

Order in unit	Item	Question Format	Content Area	Difficulty 1	Difficulty 2
1	M033q01	Multiple Choice	Space and shape	429.6	
2	M474q01	Short Response	Space and shape	444.3	
3	M155q01	Open Constructed Response	Change and relationships	486.2	
4	M155q02t	Open Constructed Response	Change and relationships	492.6	528.9
5	M155q03t	Open Constructed Response	Space and shape	629.8	698.3
6	M155q04t	Complex Multiple Choice	Space and shape	521.5	
7	M411q01	Short Response	Change and relationships	542.2	
8	M411q02	Multiple Choice	Uncertainty	567.8	
9	M803q01t	Short Response	Quantity	636.0	
10	M442q02	Closed Constructed Response	Change and relationships	592.7	
11	M462q01t	Open Constructed Response	Space and shape	677.6	
12	M034q01t	Closed Constructed Response	Space and shape	571.4	

Difficulty 1 and difficulty 2 refer to the PISA score point difficulties of each item. Partial credit items have two difficulty estimates.

Table 4: Summary of item characteristics in science block S4 (2006)

Order in unit	Item	Question Format	Process	Difficulty
1	S510Q01	Complex Multiple Choice	Explaining phenomena scientifically	536.2
2	S510Q04	Complex Multiple Choice	Explaining phenomena scientifically	600.3
3	S326Q01	Open response	Using scientific evidence	513.6
4	S326Q02	Open response	Using scientific evidence	487.4
5	S326Q03	Multiple Choice	Using scientific evidence	512.1
6	S326Q04t	Complex Multiple Choice	Explaining phenomena scientifically	689.9
7	S408Q01	Multiple Choice	Explaining phenomena scientifically	496.1
8	S408Q03	Open response	Explaining phenomena scientifically	647.6
9	S408Q04	Complex Multiple Choice	Explaining phenomena scientifically	552.2
10	S408Q05	Multiple Choice	Identifying scientific issues	594.5
11	S437Q01	Multiple Choice	Explaining phenomena scientifically	442.2
12	S437Q03	Multiple Choice	Explaining phenomena scientifically	559.5
13	S437Q04	Multiple Choice	Explaining phenomena scientifically	514.3
14	S437Q06	Open response	Explaining phenomena scientifically	421.8
15	S415Q07	Complex Multiple Choice	Identifying scientific issues	454.6
16	S415Q02	Multiple Choice	Explaining phenomena scientifically	410.2
17	S415Q08	Complex Multiple Choice	Identifying scientific issues	454.6

Table 5: Summary of item characteristics in science block S1 (2009)

Order in unit	Item	Question Format	Process	Difficulty
1	S465Q01	Open response	Using scientific evidence	724.2
2	S465Q02	Multiple Choice	Explaining phenomena scientifically	496.8
3	S465Q04t	Multiple Choice	Explaining phenomena scientifically	612.0
4	S131Q02t	Open response	Using scientific evidence	556.6
5	S131Q04t	Open response	Identifying scientific issues	568.2
6	S428Q01t	Multiple Choice	Using scientific evidence	485.9
7	S428Q03t	Multiple Choice	Using scientific evidence	428.4
8	S428Q05	Open response	Explaining phenomena scientifically	442.2
9	S514Q02	Open response	Using scientific evidence	356.2
10	S514Q03	Open response	Explaining phenomena scientifically	574.0
11	S514Q04	Complex Multiple Choice	Using scientific evidence	542.0
12	S438Q01t	Complex Multiple Choice	Identifying scientific issues	351.2
13	S438Q02	Multiple Choice	Identifying scientific issues	469.1
14	S438Q03t	Open response	Identifying scientific issues	592.3
15	S415Q07t	Complex Multiple Choice	Identifying scientific issues	454.6
16	S415Q02	Multiple Choice	Explaining phenomena scientifically	410.2
17	S415Q08t	Complex Multiple Choice	Identifying scientific issues	525.3

Tables 6, 7 and 8 show the test designs for PISA 2003, 2006 and 2009, respectively. The blocks marked in bold are those selected for analysis. It may be noted that blocks in the first position should be fully comparable across cycles, while those in other positions may not, since they tend to be preceded by blocks of varying domains. It may also be noted that booklets were assigned to students at random, so each booklet represents a random and equivalent sub-sample.

Table 6: PISA 2003 test design

Booklet	P1	P2	P3	P4
1	M1	M2	M4	R1
2	M2	M3	M5	R2
3	M3	M4	M6	PS1
4	M4	M5	M7	PS2
5	M5	M6	S1	M1
6	M6	M7	S2	M2
7	M7	S1	R1	M3
8	S1	S2	R2	M4
9	S2	R1	PS1	M5
10	R1	R2	PS2	M6
11	R2	PS1	M1	M7
12	PS1	PS2	M2	S1
13	PS2	M1	M3	S2

P1=position 1, P2=position 2, etc. M=mathematics, R=reading, S=science, PS=problem solving.

Clusters marked in bold are those selected for analysis.

Table 7: PISA 2006 test design

Booklet	P1	P2	P3	P4
1	S1	S2	S4	S7
2	S2	S3	M3	R1
3	S3	S4	M4	M1
4	S4	M3	S5	M2
5	S5	S6	S7	S3
6	S6	R2	R1	S4
7	S7	R1	M2	M4
8	M1	M2	S2	S6
9	M2	S1	S3	R2
10	M3	M4	S6	S1
11	M4	S5	R2	S2
12	R1	M1	S1	S5
13	R2	S7	M1	M3

P1=position 1, P2=position 2, etc. M=mathematics, R=reading, S=science. Clusters marked in bold are those selected for analysis.

Table 8: PISA 2009 test design

Booklet	P1	P2	P3	P4
1	M1	R1	R3A	M3
2	R1	S1	R4A	R7
3	S1	R3A	M2	S3
4	R3A	R4A	S2	R2
5	R4A	M2	R5	M1
6	R5	R6	R7	R3A
7	R6	M3	S3	R4A
8	R2	M1	S1	R6
9	M2	S2	R6	R1
10	S2	R5	M3	S1
11	M3	R7	R2	M2
12	R7	S3	M1	S2
13	S3	R2	R1	R5

P1=position 1, P2=position 2, etc. M=mathematics, R=reading, S=science. Clusters marked in bold are those selected for analysis.

Analysis Strategy

The analyses documented here consist, essentially, of comparing percent correct, missing, and not reached by cycle and position (P1 2003 = position 1, 2003, P1 2009 = position 1, 2009, P4 2003 = position 4, 2003, P4 2009 = position 1, 2009, etc.).

One would expect, due to test fatigue, that the percent correct will generally be lower and the percent missing and not reached would be higher in position 4 relative to position 1. One would also expect the response patterns for items in position 1 to be stable, all other things being equal. However, if the hypothesis about a decline in standards in Ireland is to be supported, one would expect to see a decline in percent correct and a corresponding increase in percent missing and not

reached in *both* positions. If the disengagement hypothesis is to be supported, one would expect stable percent correct and missing/not reached in position 1, but a decrease in percent correct (and increased missing responses) in position 4. The main focus of analyses is to compare positions 1 and 4, but in the case of R2 and M1 / M2, the response patterns associated with all four positions was analysed for Ireland in order to obtain detailed information across all parts of the booklets.

These analyses seek to further tease out these hypotheses by comparing Ireland with a small set of countries whose scores did not differ to those of Ireland in 2003 (2006) and have also remained stable in the intervening period. The OECD averages are also estimated to provide a broad indication of trends in response patterns internationally. If response patterns in Ireland differ to those in other countries and internationally, this would provide support for an idiosyncrasy in Ireland that may or may not be related to proficiency.

An analysis of the changes in response patterns of some reading items for sub-groups of students in Ireland was also conducted. This analysis of sub-groups (six in total – males and females with low, medium and high ESCS scores) seeks to provide insights as to whether any observed changes in response patterns are associated to a greater degree with particular sub-groups.

Then, response patterns by item type for a subset of items are examined. For this purpose, three groups of item types are compared: multiple choice and complex multiple choice; short response and closed constructed response; and open constructed response. This item categorisation is not identical to those used in Cosgrove and Moran (2011) and Cartwright (2011); however, in this report and in the other two, the categorisations make some distinction between items that require a written response and those that do not.

Data are generally unweighted since the analyses involve a small subset of the samples (e.g. in 2009, examination of item responses on one booklet equates to one-thirteenth of the sample or about 300 students). An exception is in the computation of the OECD averages, where senate weights were applied to allow OECD countries to contribute equally to the OECD average. When computing percent correct for partial credit items, the percent of partially and fully correct responses was summed. If this is not done (i.e. if partially and fully correct responses are weighted in some way), the percentages for zero, credit, missing and not reached would not sum to 100. Standard errors are not reported since the method to produce these ('bootstrapping' using the PISA replicate weights; OECD, 2009) has not been applied. Therefore, all results are indicative of broad response patterns and have not been subjected to significance tests.

In order to check whether any observed changes in response patterns could have been due to differences in pagination across positions and cycles, each of the clusters examined was checked for variations in pagination. In all domains, cycles and positions, each cluster began on an even (left-sided) page, so students were always presented with the stimulus texts and questions in the same layout. Had a cluster begun on an odd page in some instances, responses could have been affected (e.g. with different demands for page-turning), but this turned out not to be the case.

Results

Trends in Achievement in Ireland, OECD and Comparison Countries

Table 9 shows, for Ireland and three comparison countries (Belgium, the Netherlands, and New Zealand), mean reading scores for 2003 and 2009. These countries are included since (i) their scores did not differ to Ireland's in 2003 and (ii) their scores remained stable since 2003. Ireland has by far the largest change in average achievement since 2003 – a decline of around 20 score points.

Table 10 shows this information for mathematics for 2006 and 2009, this time including comparison countries Iceland, Poland and Slovenia (i.e. countries whose mathematics achievement did not differ to Ireland's in 2006 and which have remained stable 2006-2009).

Table 11 shows mean scores for 2006 and 2009 for science for Ireland along with Belgium, Hungary and the United Kingdom (again, these countries were chosen for the same reasons as making comparisons in the cases of reading and mathematics).

Table 9: Mean reading scores and standard errors – Belgium, Ireland, the Netherlands, and New Zealand – 2003 and 2009

Country	PISA 2003		PISA 2009	
	Mean	SE	Mean	SE
Belgium	507	2.6	506	2.3
Ireland	516	2.6	496	3.0
Netherlands	513	2.9	508	5.1
New Zealand	522	2.5	521	2.4

Table 10: Mean mathematics scores and standard errors – Iceland, Ireland, Poland and Slovenia – 2006 and 2009

Country	PISA 2006		PISA 2009	
	Mean	SE	Mean	SE
Iceland	506	1.8	507	1.4
Ireland	502	2.8	487	2.5
Poland	495	2.4	495	2.8
Slovenia	505	1.0	501	1.2

Table 11: Mean science scores and standard errors – Belgium, Hungary, Ireland, and the United Kingdom – 2006 and 2009

Country	PISA 2006		PISA 2009	
	Mean	SE	Mean	SE
Belgium	510	2.5	507	2.5
Hungary	504	2.7	503	3.1
Ireland	508	3.2	508	3.3
United Kingdom	515	2.3	514	2.5

Analyses of Reading Items, 2003 and 2009

Table 12 shows, for Ireland and the three comparison countries, average percent correct, incorrect (i.e. of items attempted), missing and not reached for block R2 in positions 1 and 4 in 2003 and 2009. The OECD average is also shown to illustrate the general international trends².

Table 12: Average percent correct, incorrect, missing and not reached for block R2, positions 1 and 4, 2003 and 2009 - Belgium, Ireland, the Netherlands, New Zealand, and OECD averages

% Correct	P1 2003	P1 2009	P4 2003	P4 2009
Belgium	69.7	67.2	58.3	56.2
Ireland	65.1	64.4	59.9	46.5
Netherlands	70.8	68.6	63.5	57.0
New Zealand	70.4	69.3	62.5	56.8
OECD Average	65.6	65.9	54.4	52.4
% Incorrect				
Belgium	25.3	27.0	26.4	29.0
Ireland	31.4	30.7	32.1	33.9
Netherlands	27.9	30.0	33.6	39.3
New Zealand	24.1	25.3	26.3	28.1
OECD Average	28.2	28.3	28.2	29.2
% Missing				
Belgium	5.0	5.8	8.0	8.1
Ireland	3.5	4.9	6.1	10.2
Netherlands	1.3	1.4	2.3	1.7
New Zealand	5.5	5.4	6.9	7.8
OECD Average	6.2	5.8	10.4	10.8
% Not Reached				
Belgium	0.0	0.0	7.3	6.7
Ireland	0.0	0.0	1.9	9.4
Netherlands	0.0	0.0	0.6	2.0
New Zealand	0.0	0.0	4.3	7.3
OECD Average	0.0	0.0	7.0	7.6
% Missing + Not Reached				
Belgium	5.0	5.8	15.3	14.8
Ireland	3.5	4.9	7.9	19.6
Netherlands	1.3	1.4	2.9	3.6
New Zealand	5.5	5.4	11.1	15.1
OECD Average	6.3	5.8	17.5	18.4

² The OECD average does not comprise the same set of countries across all cycles. Since PISA 2006, Chile, Estonia, Israel, the Slovak Republic and Slovenia have joined. As already noted, average percentages were computed on the pooled OECD datasets which were weighted such that each country contributes equally to the averages.

As expected, across all countries and both cycles, percent correct is lower in position 4 relative to position 1. Belgium shows the most stable pattern of percent correct across cycles and block positions, while in Ireland, the Netherlands and New Zealand, there is a marked decrease in percent correct for position 4 in 2009. There is also a small decrease, of about two percentage points, across the OECD as a whole, for these items in position 4. However, while the drop in the Netherlands and New Zealand is in the region of 6 percentage points, the drop is much larger in Ireland – 13.4 percentage points.

A comparison of the percentages correct for position 1 in 2003 and 2009 indicates that in all countries, including Ireland, as well as the OECD average, they are very similar, suggesting that on this measure at least, students were performing at about the same level in 2009 as in 2003.

Interestingly, of attempted items, percent incorrect is relatively stable in both positions, and the biggest increase (5.7%) is associated with the Netherlands for position 4. In the case of Ireland, this indicates that the decrease in percent correct is associated with increases in missing responses rather than incorrect attempts.

Turning now to the percentages of missing responses (as opposed to not reached ones), it can be seen that rates of missing responses are broadly similar across cycles in position 1 for the four countries as well as on average across the OECD. However, an examination of missing responses in position 4 indicates that Ireland shows the greatest increase, from about 6% to 10%.

Not reached items are confined to position 4 in both cycles, and across all four countries there has been an increase in the percentages of not reached items. However, once again, Ireland shows the most marked increase – from 1.9% to 9.4%. On average across the OECD, the percentage of not reached responses has increased only marginally, from 7.0% to 7.6%. Thus in 2000, students in Ireland had much lower rates of not reached responses on this reading block than on average across the OECD, yet in 2009, the rate of not reached responses exceeded the OECD average.

Adding the missing and not reached percentages (bottom portion of Table 12) gives a global measure of items that were not attempted. Focusing on position 4, it can be seen that students in the Netherlands had the lowest rates of non-attempted questions in both cycles. Rates exceed 10% in both cycles for the other three countries, but the increase in the case of Ireland (from 7.9% to 19.6%) is considerably more marked. Again, this stands in contrast to the OECD average, where there is only a slight increase (from 17.5% to 18.4%) of missing and not reached responses combined.

Table 13 shows the equivalent information as Table 12, this time for Block R1. This second block was analysed in order to assess the possibility that the large changes in response patterns in position 4 in Ireland were due to some 'rogue' feature(s) of block R2. However, the patterns for block R1 are highly similar to block R2 and Ireland again shows the most marked decline in percent correct in position 4, along with the most marked increase in missing and not reached responses in position 4.

Also consistent with the patterns associated with block R2, the percentages of correct and missing responses in position 1 in Ireland have remained stable. The only notable difference between the data shown in Tables 12 and 13 is that percent correct is higher for block R1, i.e. it comprises an easier set of items than block R2.

In contrast to block R2, however, there is a tendency for percent of attempted incorrect items to increase in position 4 in 2009 relative to 2003, except in the case of New Zealand.

Table 13: Average percent correct, incorrect, missing and not reached for block R1, positions 1 and 4, 2003 and 2009 - Belgium, Ireland, the Netherlands, New Zealand, and OECD averages

% Correct	P1 2003	P1 2009	P4 2003	P4 2009
Belgium	75.3	73.1	68.1	63.7
Ireland	71.7	70.4	68.8	58.0
Netherlands	76.0	75.7	69.8	64.7
New Zealand	72.8	75.2	68.6	67.6
OECD Average	69.9	70.3	61.2	58.3
% Incorrect				
Belgium	19.4	21.7	21.6	25.7
Ireland	24.2	24.1	23.9	28.0
Netherlands	21.7	22.8	25.9	32.0
New Zealand	23.2	20.4	21.6	22.3
OECD Average	23.5	23.6	23.7	27.3
% Missing				
Belgium	5.3	5.2	5.7	7.2
Ireland	4.1	5.0	4.9	7.8
Netherlands	2.3	1.5	3.2	2.5
New Zealand	4.0	4.2	6.5	6.3
OECD Average	6.5	6.0	8.5	9.0
% Not Reached				
Belgium	0.0	0.0	4.6	3.4
Ireland	0.0	0.5	2.4	6.2
Netherlands	0.0	0.0	1.1	0.8
New Zealand	0.0	0.2	3.3	3.8
OECD Average	0.1	0.1	6.6	5.4
% Missing + Not Reached				
Belgium	5.3	5.2	10.2	10.5
Ireland	4.1	5.5	7.3	14.0
Netherlands	2.3	1.5	4.2	3.4
New Zealand	4.0	4.4	9.8	10.0
OECD Average	6.6	6.1	15.0	14.4

Tables 14, 15 and 16 show, for block R2, average percent correct, incorrect and percent missing in 2003 and 2009 in positions 1 and 4 for multiple choice, short response, and open response items, respectively, for Ireland and the three comparison countries, along with the respective OECD averages. This is examined to see whether changes in response patterns might be more strongly associated with some item types compared to others.

Across all three item types, percent correct tends to be more stable across cycles in position 1. Also across all item types, there is a general trend of a reduction in percent correct in position 4, but the reduction in percent correct in Ireland across all item types is considerably larger than in the other three countries.

Percent incorrect (of attempted items) shows a slightly different pattern. Regardless of item type, it tends, by and large, to be stable in position 1 across the two cycles and the three item types. Patterns are less stable in position 4. Focusing on Ireland, it appears that while there has been an increase in percent of incorrect (attempted) responses in the case of multiple choice items, percent incorrect is stable for short and open response item types.

The tables also show that, irrespective of item difficulty, missing responses are lowest for multiple choice questions and highest for ones that require a longer written response. Across cycles and positions, percent missing on multiple choice items has stayed relatively stable. The same observation can be made with respect to missing responses on short response questions – with the exceptions of position 4 in Ireland and New Zealand where missing responses have increased from about 3% to 4.5%. Missing responses tend to be higher for open response items (with the exception of the Netherlands, with missing responses on these items less than 3% across item types and positions). Some small increases are apparent but these are unlikely to be of any substantial significance (e.g. 12.2% to 13.4% in position 4 in Belgium). A notable exception is position 4 in the case of Ireland, where the percentage of missing open response questions rose from 9.9% to 18.2%.

Table 14: Average percent correct, incorrect and missing and not reached for block R2 multiple choice items, positions 1 and 4, 2003 and 2009 - Belgium, Ireland, the Netherlands, and New Zealand, and OECD averages

% Correct	P1 2003	P1 2009	P4 2003	P4 2009
Belgium	75.2	73.6	65.5	63.0
Ireland	65.0	66.3	64.8	50.1
Netherlands	81.4	80.9	75.6	67.4
New Zealand	76.8	77.8	70.3	64.6
OECD Average	74.7	75.9	65.4	63.1
% Incorrect				
Belgium	23.7	25.2	28.7	30.3
Ireland	35.0	32.1	33.8	42.1
Netherlands	18.5	18.9	24.2	31.2
New Zealand	22.1	21.6	26.9	30.6
OECD Average	24.3	23.2	28.4	30.7
% Missing + Not Reached				
Belgium	1.1	1.2	5.8	6.7
Ireland	0.0	1.6	1.4	7.8
Netherlands	0.1	0.2	0.2	1.4
New Zealand	1.1	0.6	2.8	4.8
OECD Average	1.0	0.9	6.2	6.2

Table 15: Average percent correct, incorrect and missing and not reached for block R2 short response items, positions 1 and 4, 2003 and 2009 - Belgium, Ireland, the Netherlands, and New Zealand, and OECD averages

% Correct	P1 2003	P1 2009	P4 2003	P4 2009
Belgium	67.5	63.7	56.0	53.8
Ireland	65.0	60.8	57.8	47.0
Netherlands	63.8	59.0	56.8	49.3
New Zealand	65.1	63.4	55.7	52.5
OECD Average	60.7	58.8	50.0	47.1
% Incorrect				
Belgium	30.3	33.0	28.5	33.3
Ireland	33.7	36.3	35.8	35.5
Netherlands	35.8	40.6	40.4	47.0
New Zealand	32.3	33.5	33.9	31.1
OECD Average	31.2	33.8	28.3	30.5
% Missing + Not Reached				
Belgium	2.2	3.3	15.5	12.9
Ireland	1.3	2.9	6.4	17.5
Netherlands	0.4	0.4	2.8	3.7
New Zealand	2.6	3.1	10.4	16.4
OECD Average	8.1	7.4	21.7	22.4

Table 16: Average percent correct and missing and not reached for block R2 open response items, positions 1 and 4, 2003 and 2009 - Belgium, Ireland, the Netherlands, and New Zealand, and OECD averages

% Correct	P1 2003	P1 2009	P4 2003	P4 2009
Belgium	69.1	66.7	55.8	55.4
Ireland	72.0	70.8	62.6	46.6
Netherlands	68.1	66.3	61.5	53.8
New Zealand	71.0	68.9	63.1	55.1
OECD Average	62.9	63.8	50.1	48.8
% Incorrect				
Belgium	21.4	22.8	22.8	23.2
Ireland	20.6	20.9	24.1	24.6
Netherlands	29.2	30.7	33.8	41.1
New Zealand	18.5	20.9	19.7	23.8
OECD Average	29.8	29.4	31.5	31.7
% Missing + Not Reached				
Belgium	9.5	10.5	21.4	21.4
Ireland	7.4	8.3	13.3	28.8
Netherlands	2.7	3.0	4.7	5.1
New Zealand	10.5	10.2	17.2	21.1
OECD Average	7.3	6.8	18.4	19.5

So far, some general patterns across countries have been observed, and it has been shown that Ireland stands in marked contrast to the other three comparison countries when comparing percent correct, missing, and not reached for items in position 4 in 2003 and 2009. This is despite stable percent correct, missing and not reached for items in position 1 in 2003 and 2009.

The remainder of this section focuses on the Irish reading data. Some figures illustrate the data shown in previous tables for Ireland on an item-by-item basis.

Figures 1 to 4 show percent correct by item for block R2 for both cycles and all four positions. Figures 1 to 3 indicate relative stability in percent correct for positions 1, 2 and 3, and a large decline across all items in position 4 in 2009. Note that item R104Q01 was omitted from position 3 in 2009 so it is omitted from all cycles and positions in the figures.

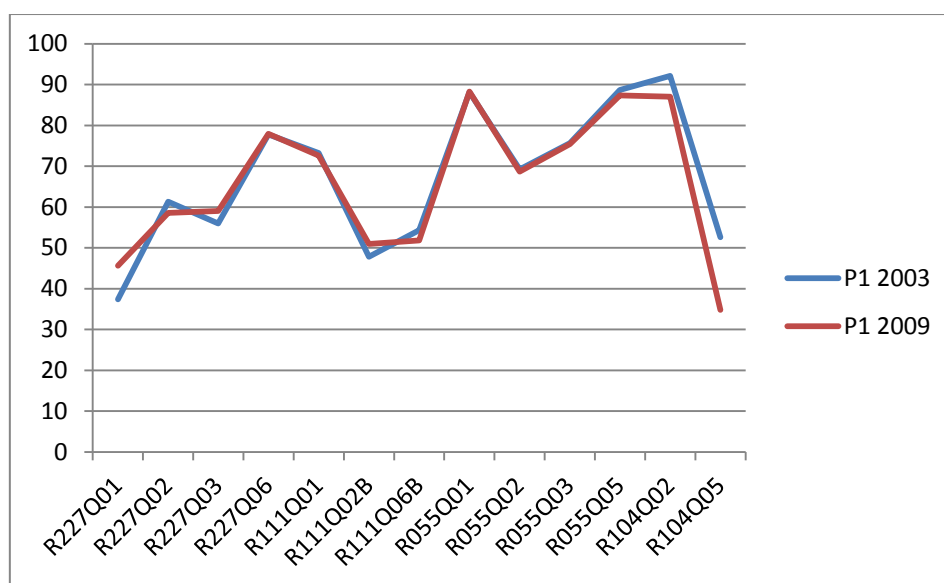


Figure 1: Percent correct by item, block R2, Ireland, 2003 and 2009, position 1

Items are shown in the order they appear in the booklet

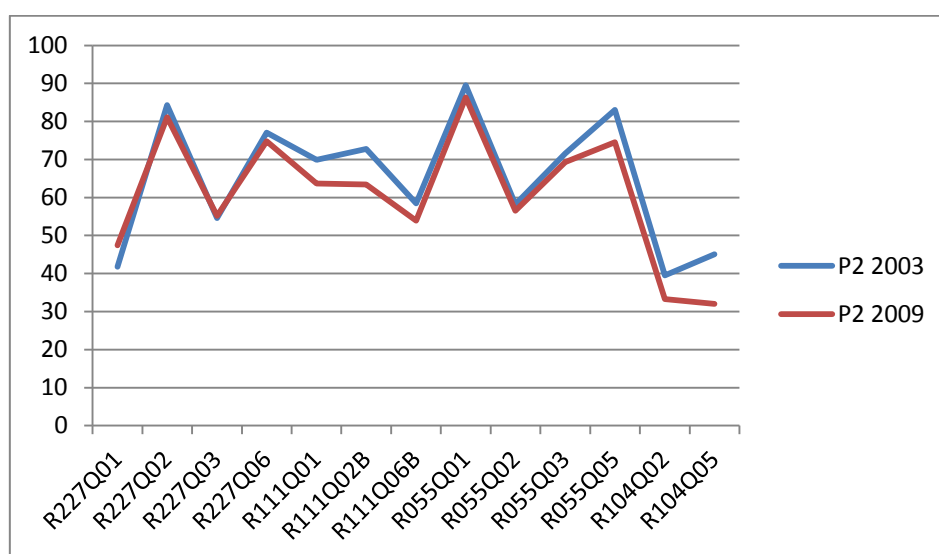


Figure 2: Percent correct by item, block R2, Ireland, 2003 and 2009, position 2

Items are shown in the order they appear in the booklet

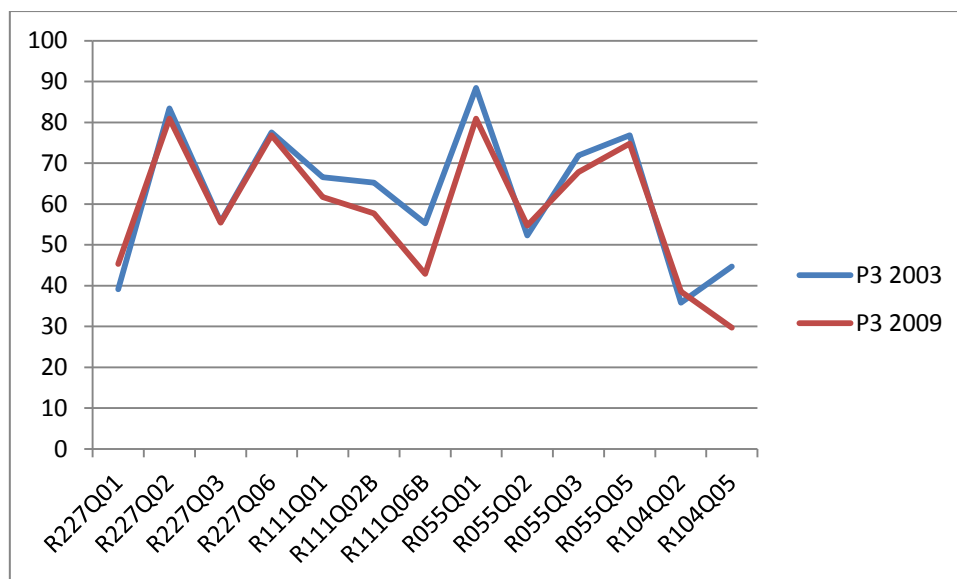


Figure 3: Percent correct by item, block R2, Ireland, 2003 and 2009, position 3
Items are shown in the order they appear in the booklet

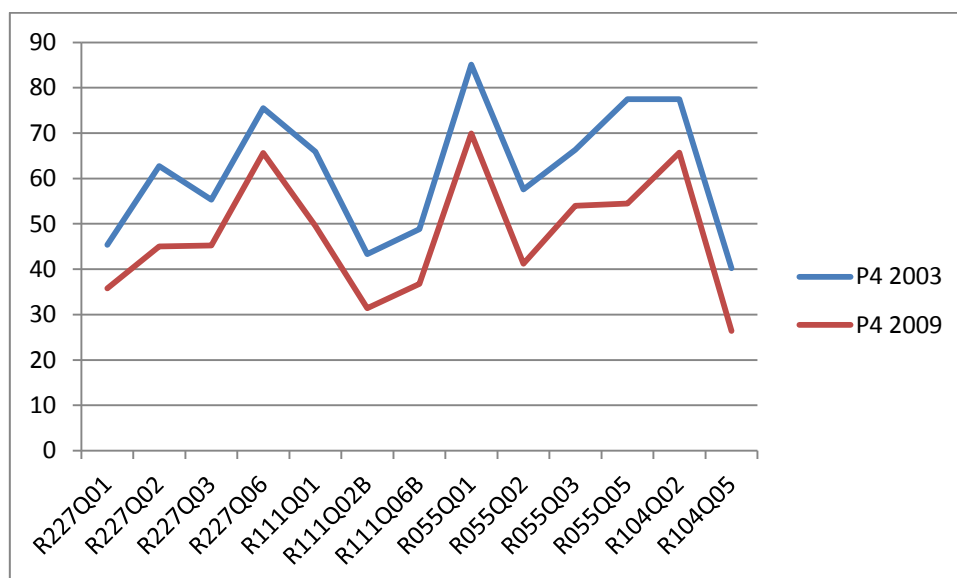


Figure 4: Percent correct by item, block R2, Ireland, 2003 and 2009, position 4
Items are shown in the order they appear in the booklet

Figures 5 to 8 show the percentage of missing plus not reached responses in Ireland in both cycles for positions 1, 2, 3 and 4, respectively. Consistent with Figure 4, Figure 8 shows a marked increase in missing and not reached responses for position 4 in 2009. Figures 5, 6 and 7 indicate some increase in missing responses for the first three positions, particularly those requiring a written response, but the increase is far more marked for position 4.

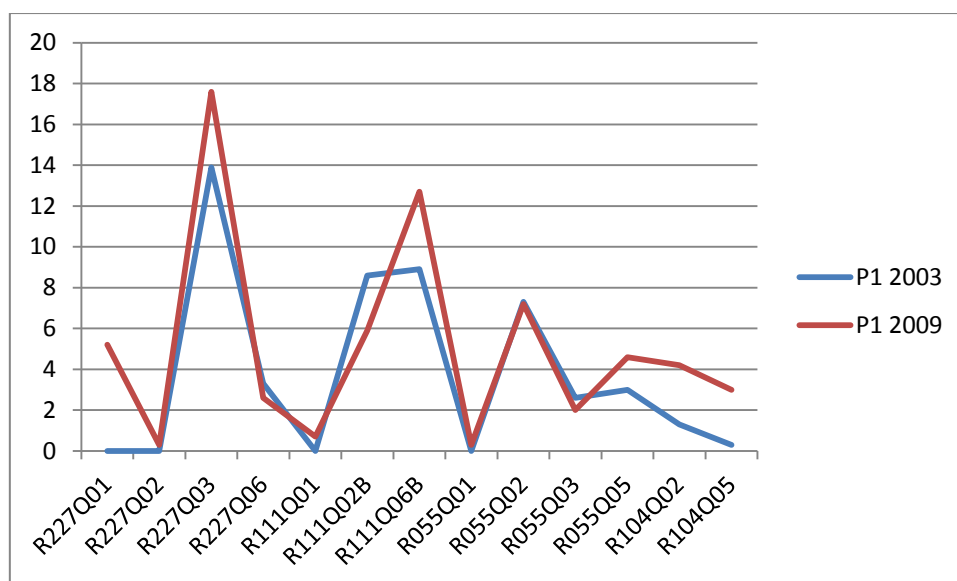


Figure 5: Percent missing plus not reached by item, block R2, Ireland, 2003 and 2009, position 1
Items are shown in the order they appear in the booklet

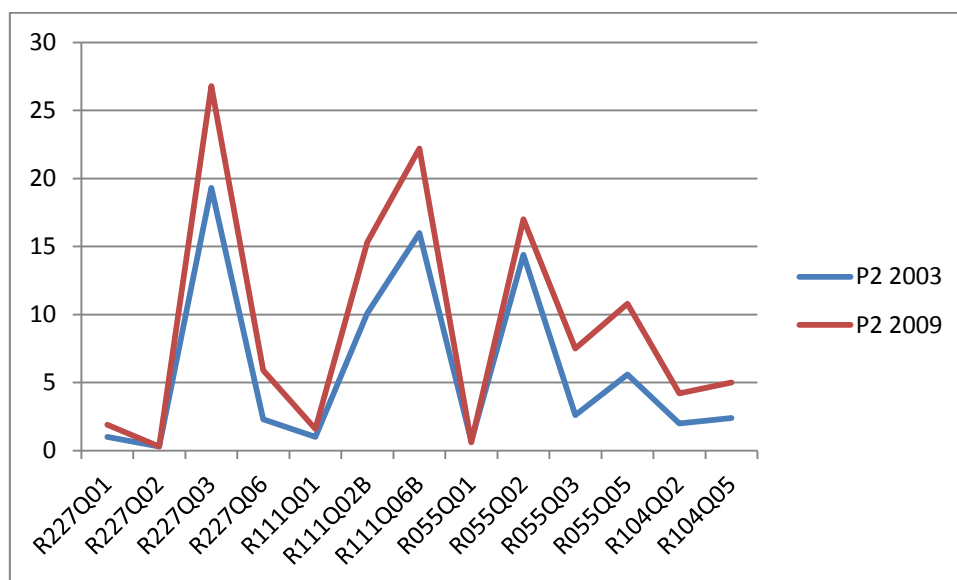


Figure 6: Percent missing plus not reached by item, block R2, Ireland, 2003 and 2009, position 2
Items are shown in the order they appear in the booklet

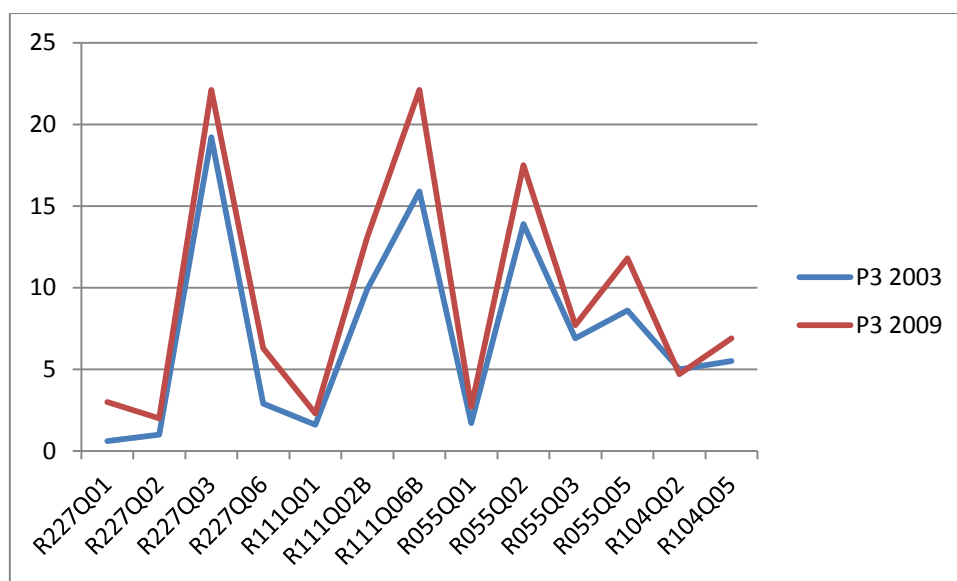


Figure 7: Percent missing plus not reached by item, block R2, Ireland, 2003 and 2009, position 3
Items are shown in the order they appear in the booklet

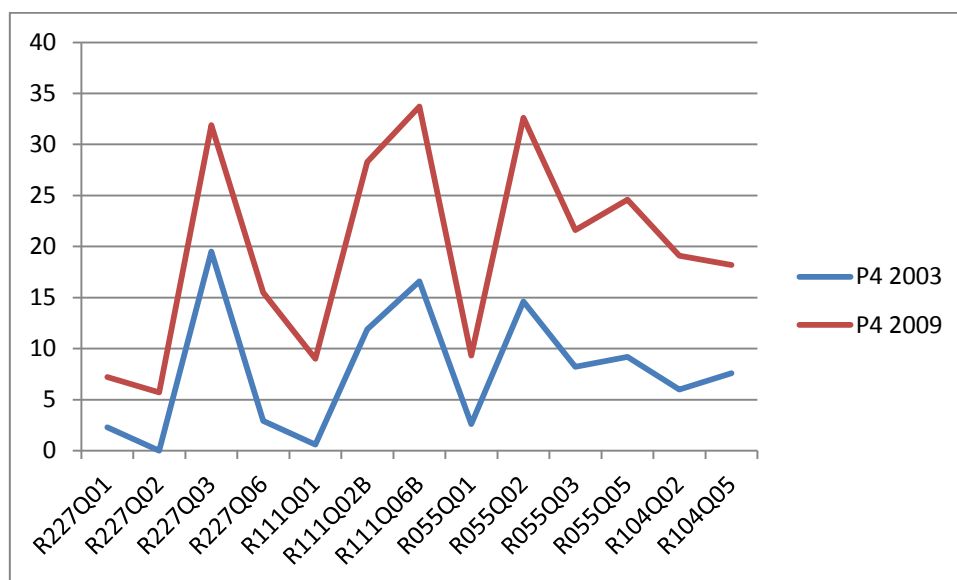


Figure 8: Percent missing plus not reached by item, block R2, Ireland, 2003 and 2009, position 4
Items are shown in the order they appear in the booklet

Figure 9 shows the percentage of not reached items in Ireland in both cycles for position 4 only. The data indicate a steady increase in not reached items in 2009 as students progress through the block – for example from about 8% halfway through this block to almost 17% at the end of the block. Close to 6% of students did not reach or attempt any items in this block in 2009. In contrast, in 2003 the rate of not reached per item was less than 2% about half way through this block and rose to just 6% – only about one-third of the equivalent figure for 2009.

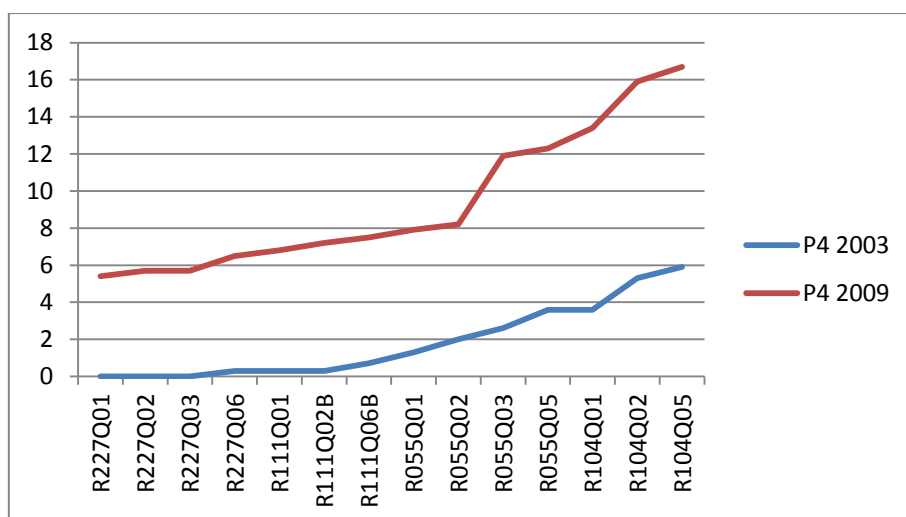


Figure 9: Percent not reached by item, block R2, Ireland, 2003 and 2009, position 4
Items are shown in the order they appear in the booklet

Figures 10 and 11 present percent correct and percent missing plus not reached for block R1 for positions 1 and 4 only. The general pattern across positions and cycles is consistent with block R2 (as shown in Figures 1, 4, 5 and 8). Figure 12 shows the percentage of not reached items for block R1 in 2003 and 2009. Again, the pattern is consistent with that found for block R2 (Figure 9). It shows that, in 2009, close to 4% of students did not attempt any items in position 4, while in 2003, this was less than 1%. The main difference between the response patterns in blocks R2 and R1 is that R2 shows a more marked increase than R1 in not reached responses in position 4.

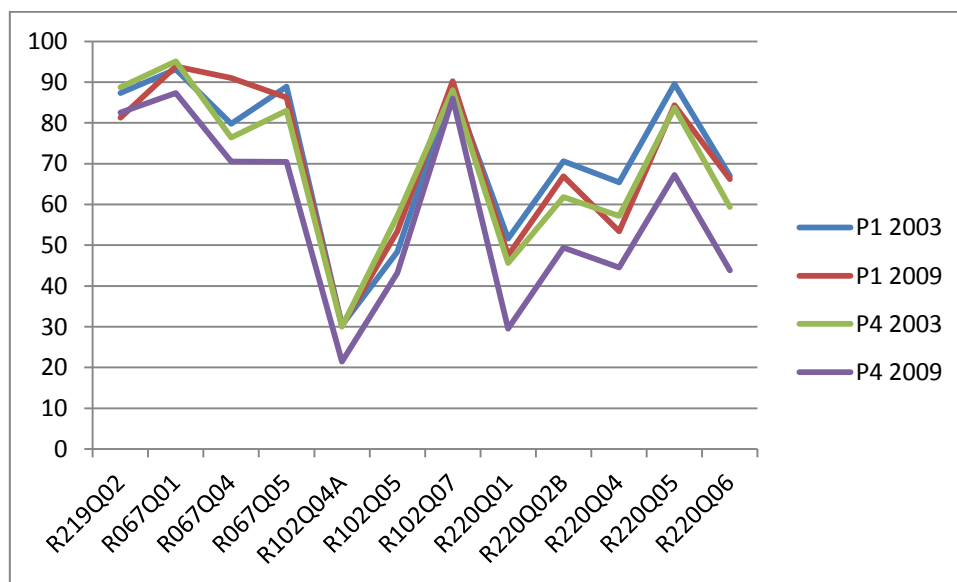


Figure 10: Percent correct by item, block R1, Ireland, 2003 and 2009, positions 1 and 4
Items are shown in the order they appear in the booklet

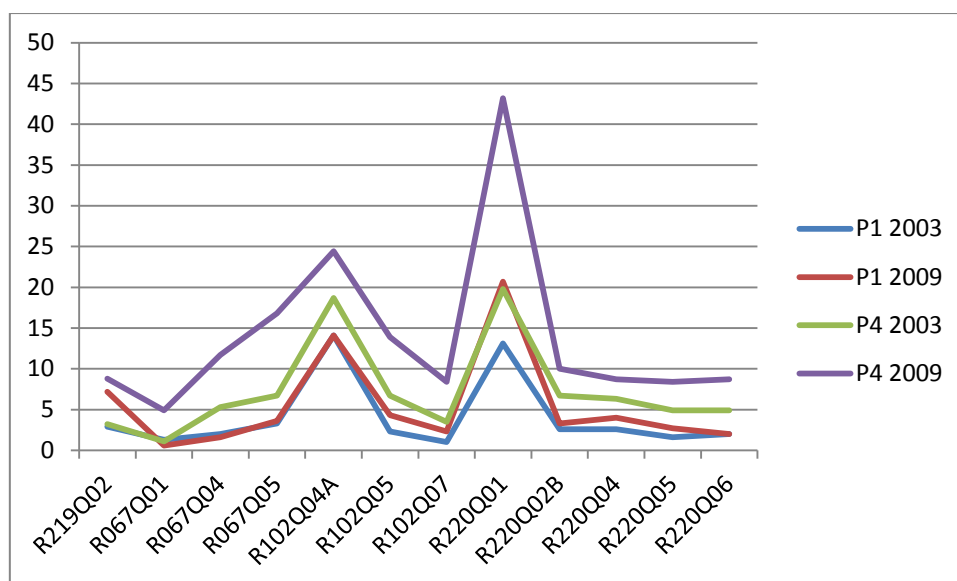


Figure 11: Percent missing plus not reached by item, block R1, Ireland, 2003 and 2009, positions 1 and 4

Items are shown in the order they appear in the booklet

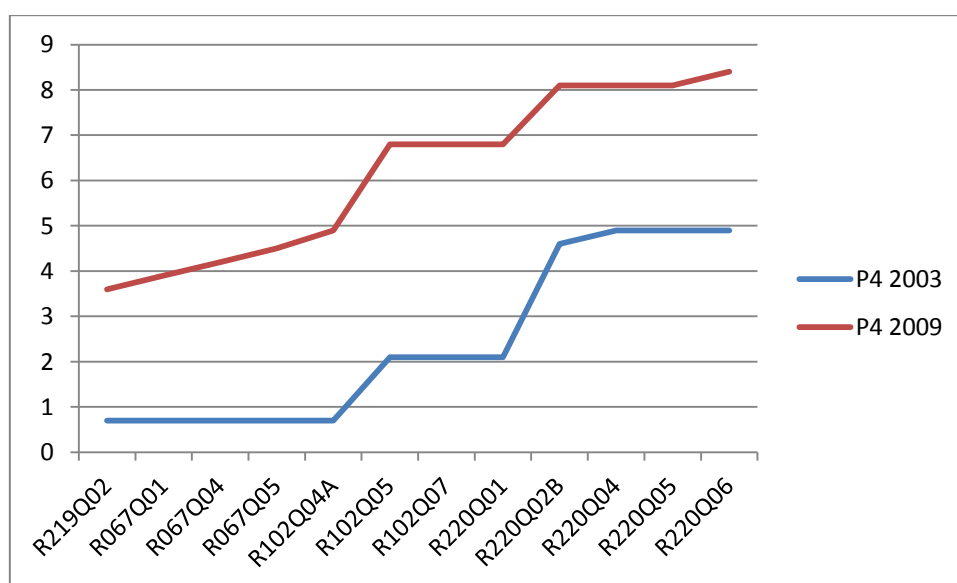


Figure 12: Percent not reached by item, block R1, Ireland, 2003 and 2009, position 4

Items are shown in the order they appear in the booklet

Trends in the Response Patterns of Sub-Groups in Reading, 2003 and 2009

One question that presents itself is whether the observed changes in response patterns described in the previous section are uniform across sub-groups of the population. Since the number of cases per booklet is small overall (about 300), it was decided to create six approximately equally-sized subgroups by gender and socioeconomic status (SES), i.e.

- Low-SES girls
- Low-SES boys
- Medium-SES girls
- Medium-SES boys
- High-SES girls
- High-SES boys³.

This results in six groups, each containing about 50 students, for 2003 and 2009.

Looking first at Table 17, it can be seen that the percentage of correct responses in position 1 in 2003 and 2009 is largely stable across the six sub-groups. However, there are two exceptions. Medium- and high-SES boys show a decrease in percent correct, by 6.3% and 7.8%, respectively. In position 4, the majority of the six groups show a decline, ranging from 10.9% (medium-SES girls) to 17.9% (low-SES boys).

Table 17: Average percent correct for block R2, positions 1 and 4, 2003 and 2009 – Irish students grouped by gender and SES

% Correct	P1 2003	P1 2009	P4 2003	P4 2009
Low-SES girls	61.6	63.5	61.0	43.1
Low-SES boys	55.7	57.3	49.5	37.0
Medium-SES girls	75.5	72.2	65.3	54.4
Medium-SES boys	69.5	63.2	64.1	49.9
High-SES girls	78.2	79.2	78.1	65.7
High-SES boys	77.2	69.4	67.5	54.1

Table 18, position 1 percent missing, indicates a general increase in the number of missing responses, and this is more pronounced among low-SES boys, medium-SES boys and girls, and high-SES boys. Table 19 also shows an increase in the number of not reached items, and this is particularly marked among low- and high-SES boys, and also low-SES girls.

Table 18: Average percent missing for block R2, positions 1 and 4, 2003 and 2009 – Irish students grouped by gender and SES

% Missing	P1 2003	P1 2009	P4 2003	P4 2009
Low-SES girls	2.8	5.5	5.4	14.7
Low-SES boys	8.3	8.6	12.8	17.6
Medium-SES girls	1.0	3.4	3.9	4.7
Medium-SES boys	3.5	6.2	5.2	10.9
High-SES girls	2.0	2.8	3.6	5.8
High-SES boys	2.3	4.2	5.0	7.2

³ The SES groupings were established by taking the 33rd and 67th percentile points on the ESCS scale on the entire weighted datasets for PISA 2003 and 2009 and using these as the cutpoints. The cutpoints are cycle-specific (i.e. the cutpoints used in 2003 were not applied to 2009 and vice versa).

Table 19: Average percent not reached for block R2, positions 1 and 4, 2003 and 2009 – Irish students grouped by gender and SES

% Not Reached	P1 2003	P1 2009	P4 2003	P4 2009
Low-SES girls	0.0	0.0	0.0	7.2
Low-SES boys	0.0	0.0	1.5	17.1
Medium-SES girls	0.0	0.0	3.2	4.7
Medium-SES boys	0.0	0.0	2.3	5.5
High-SES girls	0.0	0.0	2.4	2.9
High-SES boys	0.0	0.0	1.8	10.2

It can be concluded from Tables 17, 18 and 19 that the decline in achievement is not uniform and appears to be mediated by both socioeconomic background and gender. It can also be seen that response patterns are most stable for medium- and high-SES girls.

Table 20 shows the percentage of missing plus not reached items in each cycle and each position for the six groups. Focusing on position 4, it can be seen, consistent with the overall average shown in Table 3 for Ireland, that there has been an increase from 2003 to 2009. However, this increase is again non-uniform across groups. For example, medium- and high-SES girls show only modest increases, while the sharpest increases are associated with low-SES boys (from 14.3% to 34.7%) and low-SES girls (from 5.4% to 22.0%). The increase in non-attempted items in high-SES boys (from 6.7% to 17.4%) is also substantial.

Table 20: Average percent missing + not reached for block R2, positions 1 and 4, 2003 and 2009 – Irish students grouped by gender and SES

% Missing + Not Reached	P1 2003	P1 2009	P4 2003	P4 2009
Low-SES girls	2.8	5.5	5.4	22.0
Low-SES boys	8.3	8.6	14.3	34.7
Medium-SES girls	1.0	3.4	7.1	9.4
Medium-SES boys	3.5	6.2	7.5	16.4
High-SES girls	2.0	2.8	6.0	8.6
High-SES boys	2.3	4.2	6.7	17.4

The analyses shown in Tables 18 to 20 above were extended to examine changes in mean scores and standard deviations by the six subgroups, for all PISA domains. Figure 13 shows the mean scores of students by gender and SES group for reading in 2003, 2006 and 2009. The declines are largest (comparing 2003 and 2009) for low- and medium-SES boys (at 25 and 29 points, respectively), and the declines range from 5 points (low-SES girls) to 18 points (high-SES boys) for the remaining groups. Figure 14 shows the distribution of achievement (i.e. standard deviations) for these six groups for reading in 2003, 2006 and 2009. Overall, there has been an increase in the dispersion of achievement, and this is more marked in the case of medium-SES boys and girls, as well as for low-SES boys.

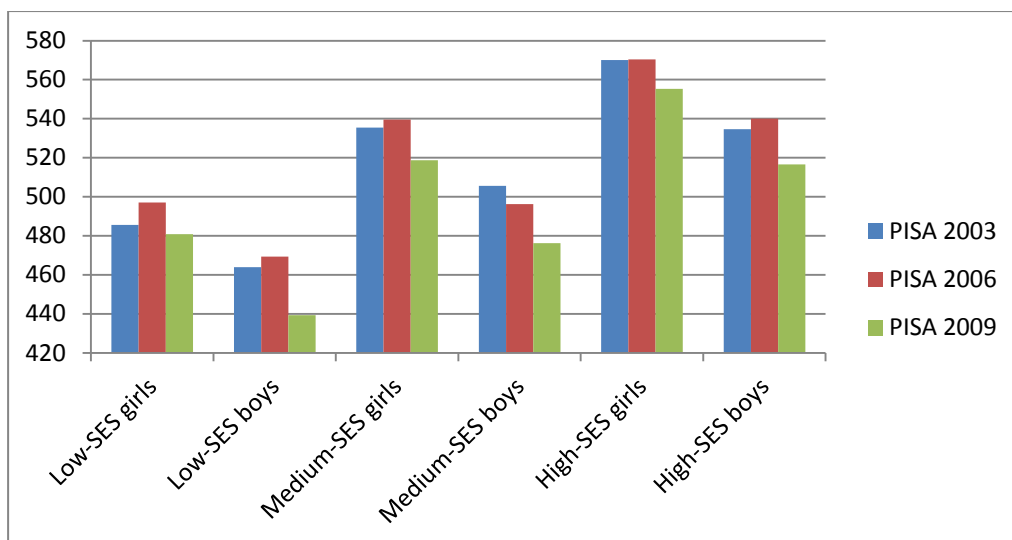


Figure 13: Mean reading scores by gender and SES group, PISA 2003, 2006 and 2009

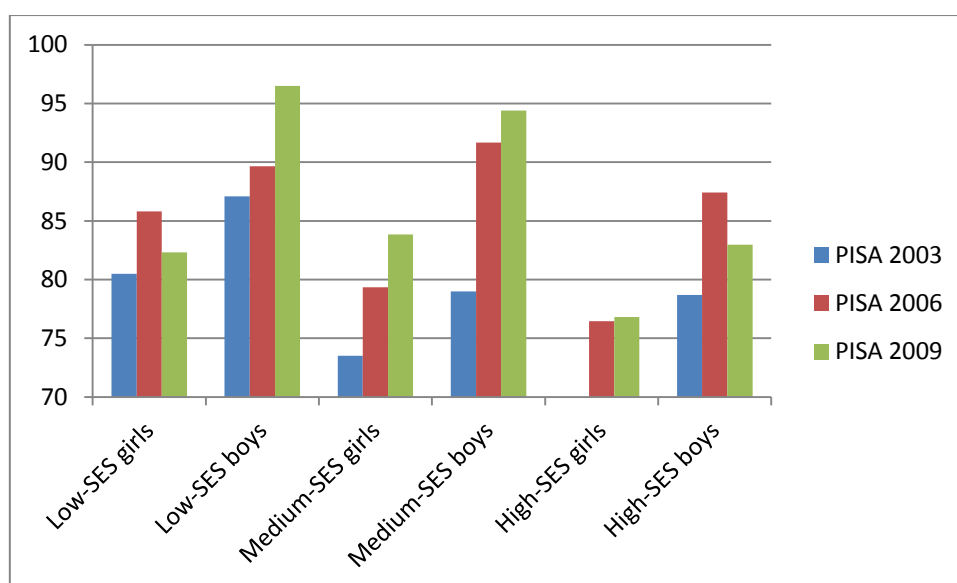


Figure 14: Standard deviations of reading scores by gender and SES group, PISA 2003, 2006 and 2009

In contrast, the pattern is not the same when one examines the mean scores and standard deviations for mathematics for these groups of students (Figures 15 and 16). There has been a relatively uniform decline in mathematics achievement (with the exception of low-SES girls) and the distribution of achievement has increased only slightly apart from high-SES boys and girls, where there has been a small decrease in the standard deviation. An exception is the distribution of achievement for low-SES boys which, like reading, has shown a large increase since 2003.

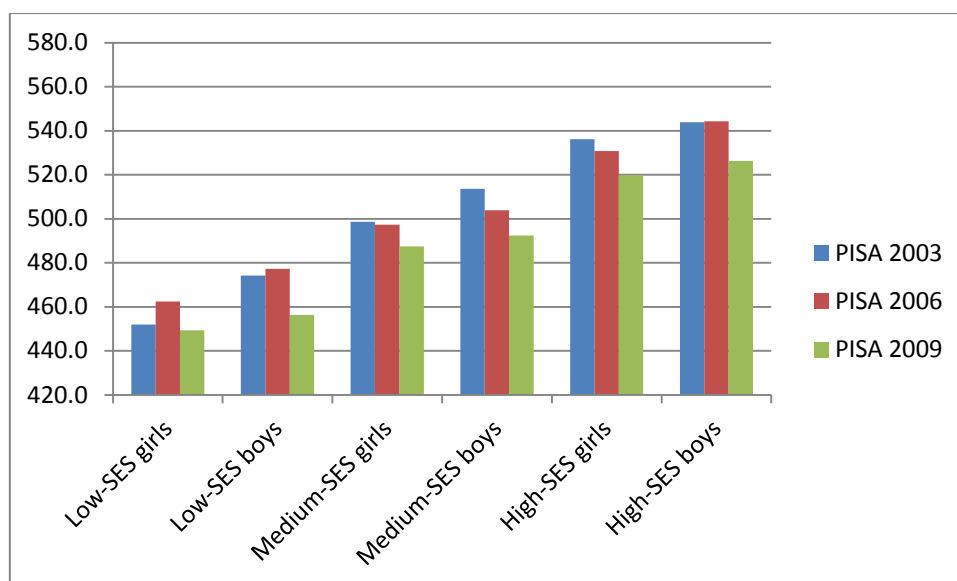


Figure 15: Mean mathematics scores by gender and SES group, PISA 2003, 2006 and 2009

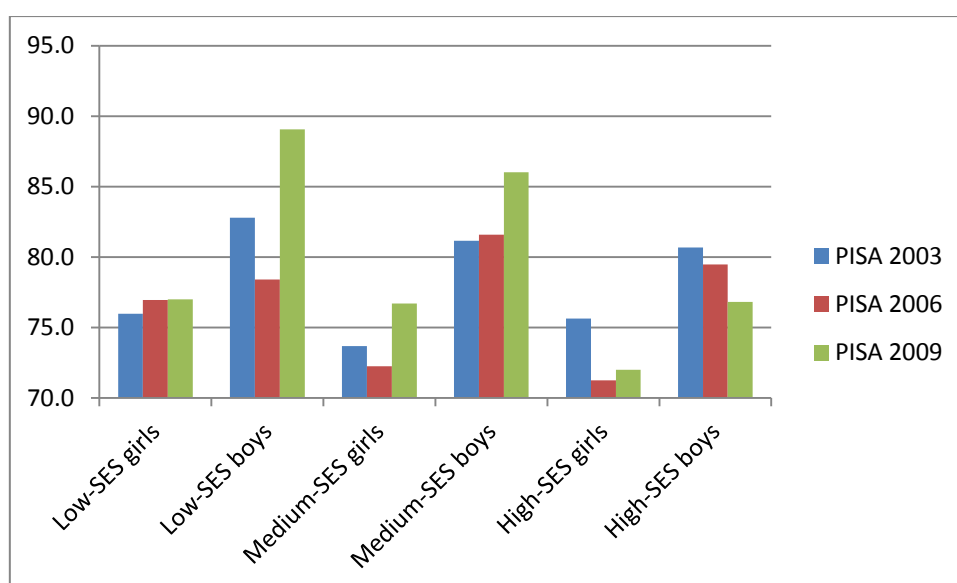


Figure 16: Standard deviations of mathematics scores by gender and SES group, PISA 2003, 2006 and 2009

In contrast to reading and mathematics, science scores have remained relatively stable across PISA cycles and gender/SES groups (Figure 17). The increase in the mean score of low-SES girls, however, is noteworthy. Also, as Figure 18 shows, there has tended to be an increase in the standard deviation of science scores, most notably among low-and medium-SES boys.

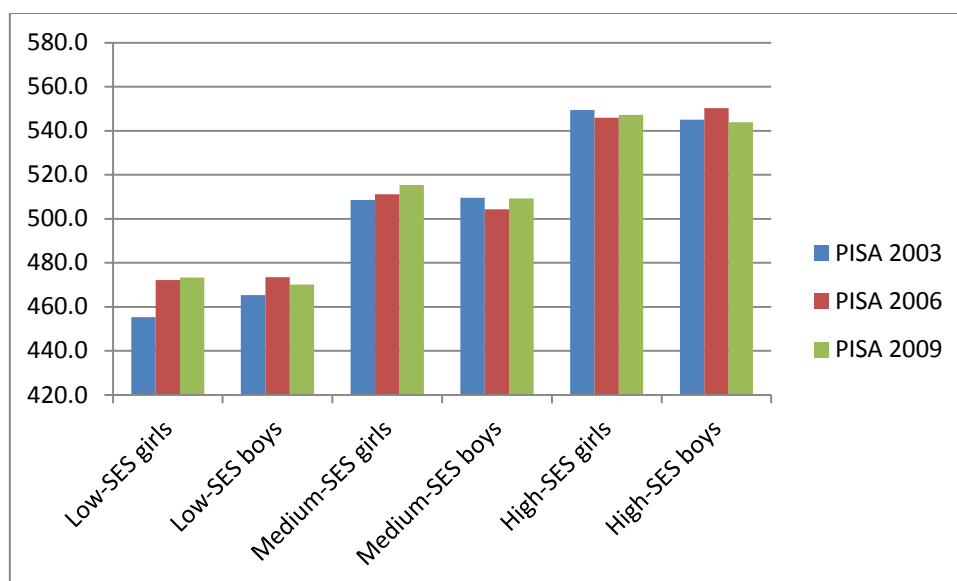


Figure 17: Mean science scores by gender and SES group, PISA 2003, 2006 and 2009

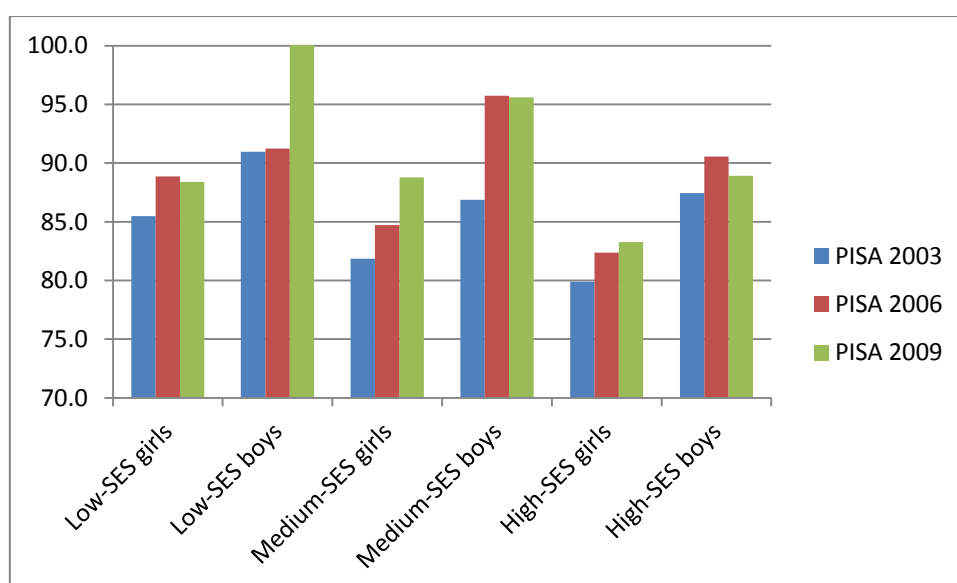


Figure 18: Standard deviations of science scores by gender and SES group, PISA 2003, 2006 and 2009

One possibility for observed changes in these six sub-groups is that there have been changes in their composition. Table 21 shows that the general increase in students speaking a language other than English or Irish is not evenly distributed across the six groups. Across SES groups, the percentage of other language speakers is higher among males than females in 2009. This represents a shift in demographics compared to 2003 and 2006.

Table 21: Percent of other language speakers, by gender and SES group

Group	2003	2006	2009
Low-SES girls	0.7	2.1	2.7
Low-SES boys	0.4	2.1	5.2
Medium-SES girls	0.7	0.8	2.4
Medium-SES boys	0.4	2.1	5.6
High-SES girls	1.1	3.1	2.0
High-SES boys	1.5	2.0	4.0

Are the Response Patterns for Link Reading Items the same as for New Reading Items?

It is possible that the response patterns associated with the link reading items differ to those of new reading items. To explore this possibility, a reading cluster introduced in 2009 – R3A – was analysed and compared with the response patterns associated with R2 in each position. Table 22 shows the average percent correct, incorrect, and missing plus not reached for the new and link clusters for each of the four booklet positions (Ireland only). There is evidence for a positioning effect when one compares positions 1 and 2, and positions 3 and 4. However, the positioning effect is much more marked for the cluster of link items compared with the new cluster. For example, percent correct for R2 declines from 66% in position 1 to 48% in position 4, while the equivalent percentages for R3A are 55% and 47%, respectively. It should be noted that there are four additional new reading clusters in the PISA 2009 test that have not yet been examined. It is possible that there are differential positioning effects associated with each of the five new reading clusters, but, nonetheless, the differences in the patterns of percent correct and missing/not reached are quite marked in Table 22.

Table 22: Percent correct, incorrect and missing plus not reached, positions 1 to 4, clusters R2 and R3A, Ireland, 2009

Position	Correct		Incorrect		Missing and not reached	
	Link cluster (R2)	New cluster (R3A)	Link cluster (R2)	New cluster (R3A)	Link cluster (R2)	New cluster (R3A)
P1	66.0	54.9	28.9	39.7	5.1	5.3
P2	60.9	52.4	29.9	41.2	9.2	6.5
P3	59.0	52.5	31.6	38.4	9.4	9.1
P4	47.8	47.1	32.5	40.8	19.7	12.1

Analyses of Mathematics Items, 2003 and 2009

Table 23 shows percentages correct, incorrect, missing and not reached for the same mathematics block (M1/M2) administered in 2006 and 2009 for Iceland, Ireland, Poland and Slovenia. Looking first at position 1 in 2006 and 2009, a small decrease in percent correct (in the region of 3 percentage points) is evident in Iceland and Ireland. With respect to position 4, percent correct is very stable across cycles, with the exception of Ireland, where the percentage of correct responses changes from 49.2% to 44.5%. This decrease, however, is not as dramatic as was observed in the case of reading and noted in the previous sections.

Table 23: Average percent correct, incorrect, missing and not reached for block M1, positions 1 and 4, 2006 and 2009 - Iceland, Ireland, Poland, and Slovenia, and OECD averages

% Correct	P1 2006	P1 2009	P4 2006	P4 2009
Iceland	57.6	54.4	46.3	48.5
Ireland	52.7	49.5	49.2	44.5
Poland	49.5	50.3	46.1	47.3
Slovenia	48.4	47.3	42.6	41.1
OECD Average	51.1	51.5	45.7	46.1
<hr/>				
% Incorrect				
Iceland	36.5	37.9	39.0	34.5
Ireland	40.4	41.5	39.4	36.9
Poland	39.2	40.0	37.0	38.3
Slovenia	38.9	39.4	37.4	39.7
OECD Average	38.8	38.9	37.9	37.4
<hr/>				
% Missing				
Iceland	5.9	7.7	10.5	10.6
Ireland	6.9	8.7	8.9	12.0
Poland	11.3	9.7	13.5	11.6
Slovenia	12.7	13.3	17.0	17.1
OECD Average	10.1	9.6	12.2	12.2
<hr/>				
% Not Reached				
Iceland	0.0	0.0	4.2	6.4
Ireland	0.0	0.3	2.5	6.6
Poland	0.0	0.0	3.4	2.8
Slovenia	0.0	0.0	3.0	2.1
OECD Average	0.0	0.0	4.2	4.3
<hr/>				
% Missing + Not Reached				
Iceland	5.9	7.7	14.6	17.0
Ireland	6.9	8.9	11.4	18.6
Poland	11.3	9.7	16.9	14.3
Slovenia	12.7	13.3	20.0	19.1
OECD Average	10.1	9.6	16.4	16.5

Comparing the percentage of (attempted) incorrect items, these are again quite stable in position 1, across countries and cycles, and in position 4 there are small decreases in the percentage of incorrect responses in the cases of Iceland and Ireland. In other words, the decrease in percent correct in position 4 is due to an increase in missing responses rather than a decrease in the percentage of attempted correct responses. The percentage of missing responses tends to be quite stable across cycles. However, in position 4, there is a small decrease in missing responses in the case of Poland (from 13.5% to 11.6%) and an increase in the case of Ireland (from 8.9% to 12.0%). Ireland also stands out in terms of the change in the percentage of not reached responses – from 2.5% in position 4 in 2006 to 6.6% in 2009. Also, combining missing and not reached responses,

across the OECD on average, about 16% of students did not attempt or did not reach position 4 in 2006 and 2009, and in Ireland the increase across cycles is more marked – from 11.4% to 18.6%.

Tables 24 to 26 show the percentages of correct, missing and not reached responses by item type for the same countries shown in Table 23.

Looking at Table 24, one can observe a 2% drop for Ireland in percent correct for M1 multiple choice items in position 1. This is the largest of the declines in percent correct in this position. In position 4, there is a 7.5% drop in percent correct in the case of Ireland. On all other cases, percent correct is quite stable with the exception of Slovenia, where it also decreased in position 4 (by just under 6%). Looking at percent missing/not reached for position 4 in particular, Ireland shows the most marked increase on this indicator for multiple choice items, from 3.2% in 2006 to 9.2% in 2009. No other country shows an increase of this magnitude. By and large, percent of (attempted) incorrect items is stable across cycles in position 1. In position 4, there is some fluctuation, e.g. a slight decrease in the case of Iceland, a slight increase in Ireland, and a somewhat larger increase in Slovenia. So, for example, in the case of Slovenia, more students that attempted M1 multiple choice items in position 4 got an incorrect answer in 2009 compared with 2006.

Table 24: Average percent correct and missing and not reached for block M1 multiple choice items, positions 1 and 4, 2006 and 2009 – Iceland, Ireland, Poland and Slovenia, and OECD averages

% Correct	P1 2006	P1 2009	P4 2006	P4 2009
Iceland	66.6	65.4	57.8	60.5
Ireland	58.7	56.6	60.9	53.4
Poland	58.8	57.4	54.9	55.8
Slovenia	60.5	60.7	58.1	52.4
OECD Average	60.2	60.5	56.3	56.5
<hr/>				
% Incorrect				
Iceland	30.7	29.9	34.8	30.2
Ireland	38.8	39.3	35.9	37.4
Poland	38.6	40.2	40.7	40.4
Slovenia	36.4	36.4	36.5	42.4
OECD Average	36.8	36.1	37.6	36.6
<hr/>				
% Missing + Not Reached				
Iceland	2.7	4.7	7.4	9.3
Ireland	2.5	4.1	3.2	9.2
Poland	2.6	2.4	4.4	3.8
Slovenia	3.1	2.9	5.4	5.2
OECD Average	3.0	3.4	6.1	6.9

Table 25 shows the percent correct, missing and not reached for short response items in the mathematics block examined. Percent correct on short response mathematics items has decreased in position 1 in all four countries considered in the table (along with an increase in incorrect responses) and Ireland does not stand out in this regard. However, Ireland's decrease in percent correct in position 4 is the largest of the four countries. Similar to Iceland, the percent of

(attempted) incorrect short response items has *decreased* in Ireland in position 4, i.e. students who did attempt short response items in position 4 were somewhat more likely to have a correct response in 2009 relative to 2006. Particularly with respect to position 4, the increase in missing responses is most marked in the case of Ireland.

Table 25: Average percent correct and missing and not reached for block M1 short response items, positions 1 and 4, 2006 and 2009 - Iceland, Ireland, Poland and Slovenia, and OECD averages

% Correct	P1 2006	P1 2009	P4 2006	P4 2009
Iceland	59.3	50.2	45.0	44.4
Ireland	51.4	45.1	43.9	39.2
Poland	49.1	39.5	43.3	44.8
Slovenia	45.4	42.9	38.2	38.0
OECD Average	50.4	49.9	43.4	43.2
% Incorrect				
Iceland	38.2	45.2	42.4	40.6
Ireland	45.1	50.1	47.7	44.5
Poland	46.5	57.9	43.8	44.1
Slovenia	47.8	49.7	47.5	48.8
OECD Average	44.3	45.3	43.1	43.5
% Missing + Not Reached				
Iceland	2.5	4.6	12.6	15.0
Ireland	3.5	4.8	8.4	16.3
Poland	4.4	2.6	12.9	11.1
Slovenia	6.8	7.4	14.3	13.2
OECD Average	5.3	4.8	13.5	13.3

Table 26 indicates that the percent correct for open response items in position 1 is quite stable. In contrast, in position 4, Ireland is the only country among those examined to show a drop in percent correct. Along with this, Ireland has the largest increase in percent missing/not reached in position 4. There is no strong discernible pattern with respect to the percentage of incorrect (attempted) responses in either position.

Table 26: Average percent correct and missing and not reached for block M1 open response items, positions 1 and 4, 2006 and 2009 - Iceland, Ireland, Poland and Slovenia, and OECD averages

% Correct	P1 2006	P1 2009	P4 2006	P4 2009
Iceland	48.6	51.5	39.2	44.6
Ireland	49.8	49.8	47.2	44.5
Poland	42.9	45.9	43.1	44.0
Slovenia	43.2	42.6	36.4	36.3
OECD Average	45.3	46.7	40.7	41.8
% Incorrect				
Iceland	38.8	34.7	38.3	30.2
Ireland	35.9	32.6	31.5	27.1
Poland	30.6	32.4	25.6	29.7
Slovenia	29.5	28.7	25.4	26.7
OECD Average	33.4	33.0	31.5	30.6
% Missing + Not Reached				
Iceland	12.6	13.8	22.5	25.2
Ireland	14.3	17.6	21.3	28.4
Poland	26.5	21.7	31.3	26.3
Slovenia	27.3	28.7	38.2	37.0
OECD Average	21.3	20.3	27.8	27.6

Figures 19 to 22 show, for Ireland, the percentages of correct responses by item, position and cycle, for block M1. Generally, percent correct is more stable in all positions for mathematics when compared to reading (as discussed previously), and is most stable in position 3. In the other three positions, particularly position 4, some items show a decline in percent correct.

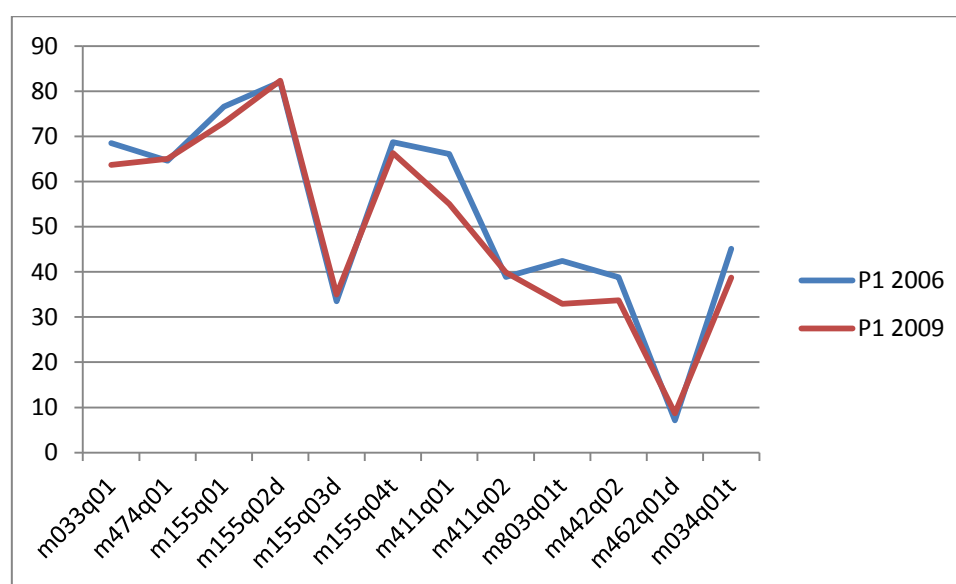


Figure 19: Percent correct by item, block M1 (M2), Ireland, 2006 and 2009, position 1

Items are shown in the order they appear in the booklet

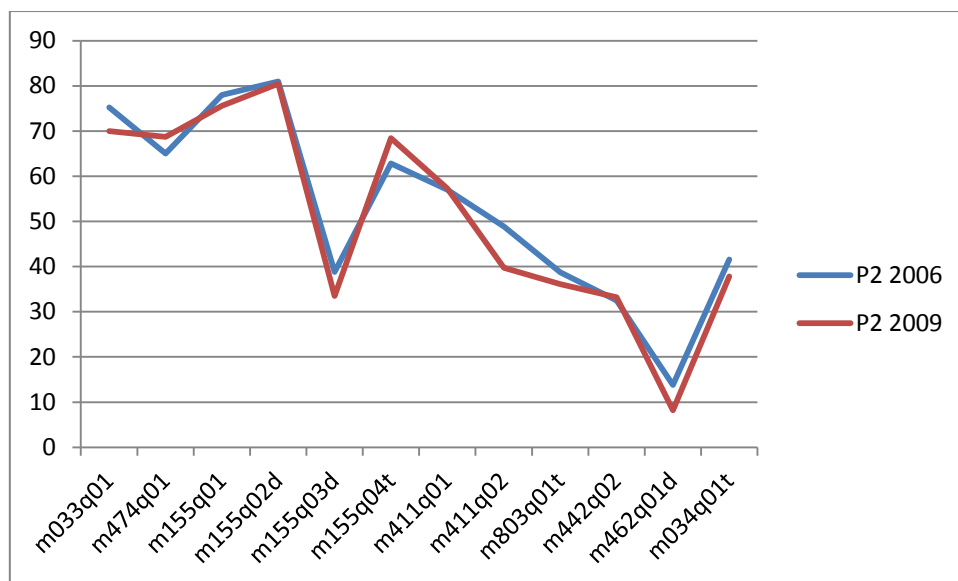


Figure 20: Percent correct by item, block M1 (M2), Ireland, 2006 and 2009, position 2
Items are shown in the order they appear in the booklet

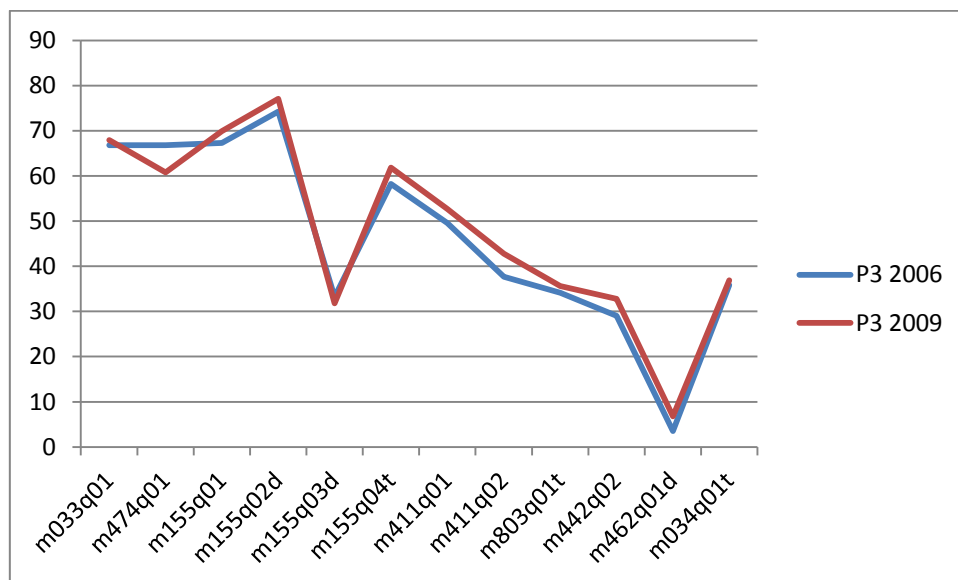


Figure 21: Percent correct by item, block M1 (M2), Ireland, 2006 and 2009, position 3
Items are shown in the order they appear in the booklet

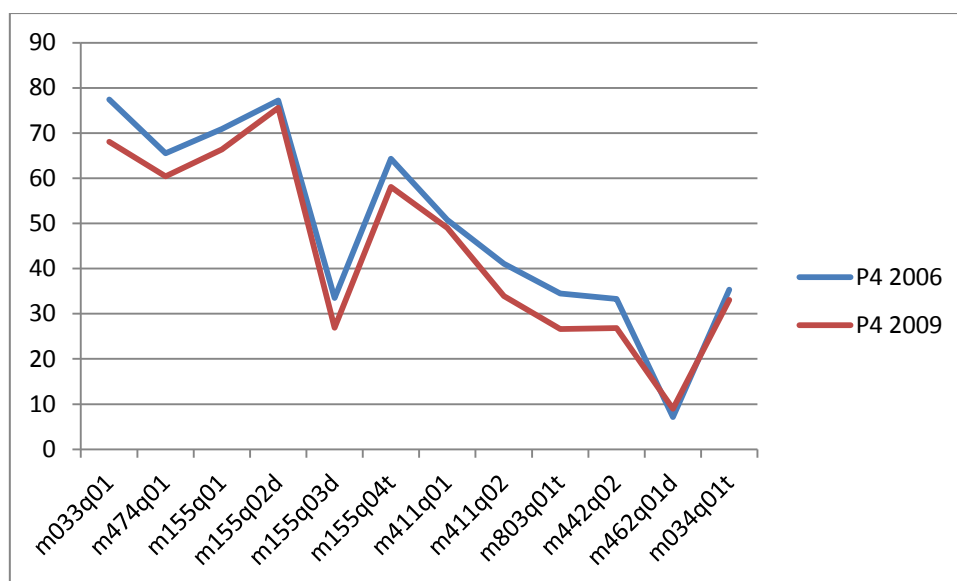


Figure 22: Percent correct by item, block M1 (M2), Ireland, 2006 and 2009, position 4
Items are shown in the order they appear in the booklet

Figures 23 to 26 show the percent of missing plus not reached items for positions 1, 2, 3 and 4, respectively, for Ireland in 2006 and 2009. In positions 1 and 2, three items show a relative increase in missing/not reached responses (two of these are written response items), while in position 3, the pattern of missing/not reached responses is very stable across cycles. Position 4 shows the greatest increase in missing/not reached responses in 2009 relative to 2006. Figure 27 shows a uniform increase in not reached responses in 2009 relative to 2006.

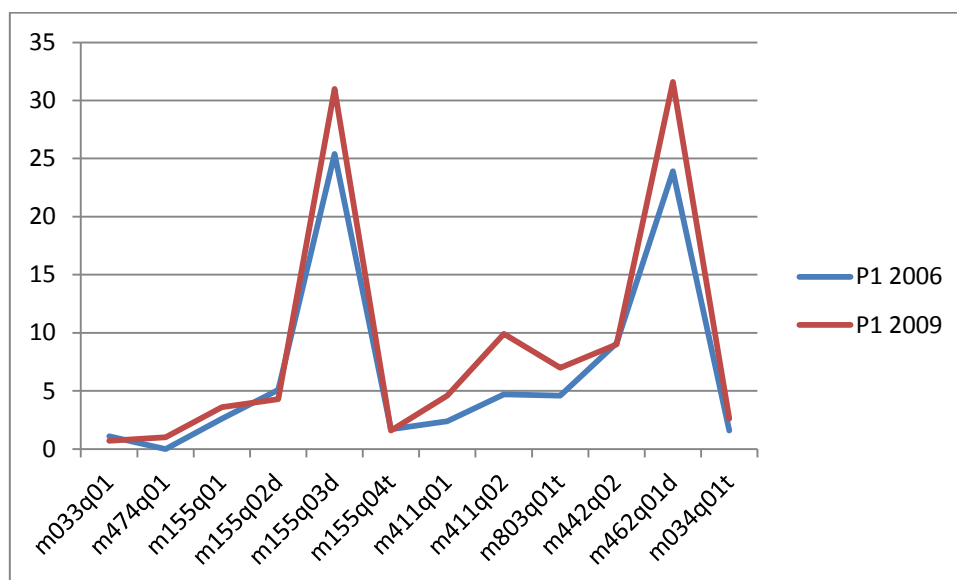


Figure 23: Percent missing plus not reached by item, block M1 (M2), Ireland, 2006 and 2009, position 1
Items are shown in the order they appear in the booklet

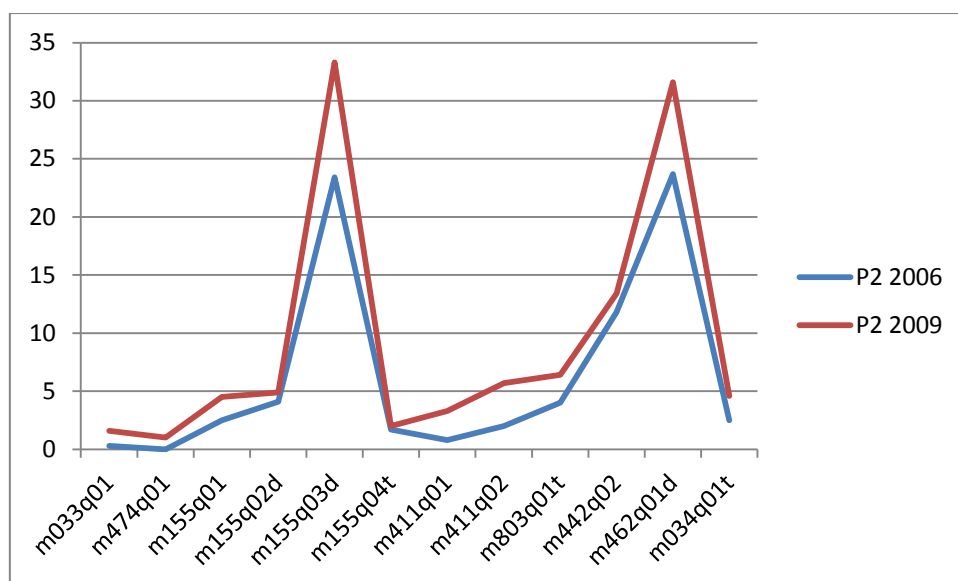


Figure 24: Percent missing plus not reached by item, block M1 (M2), Ireland, 2006 and 2009, position 2

Items are shown in the order they appear in the booklet

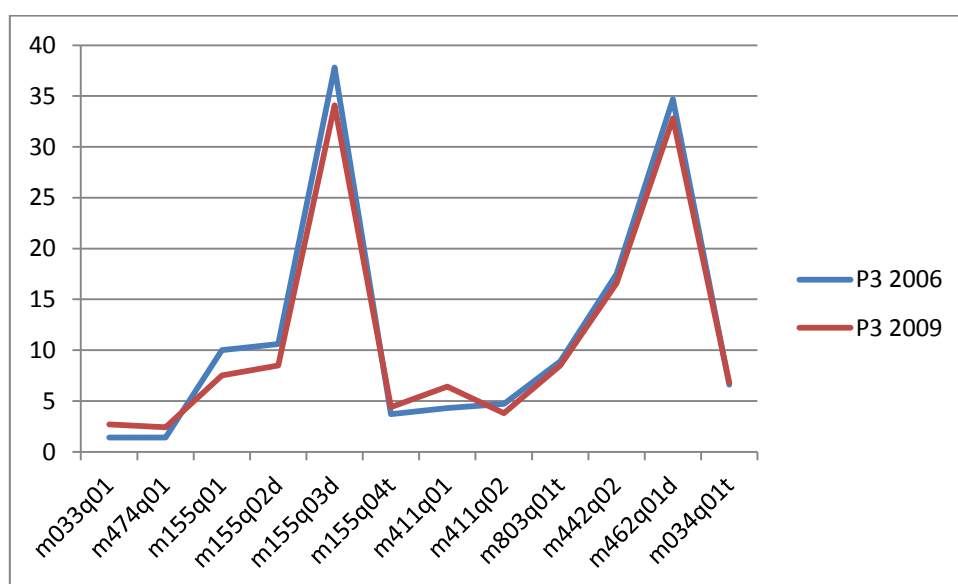


Figure 25: Percent missing plus not reached by item, block M1 (M2), Ireland, 2006 and 2009, position 3

Items are shown in the order they appear in the booklet

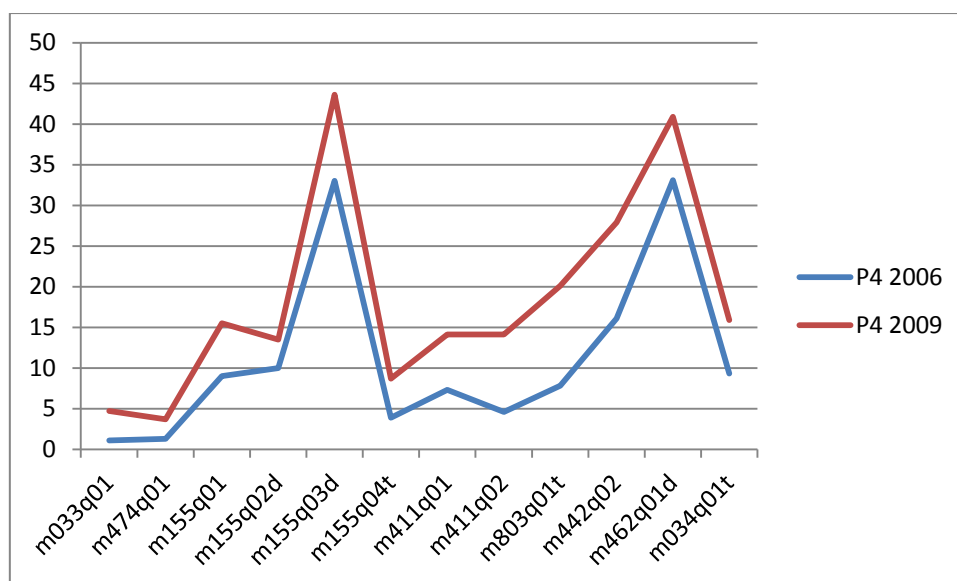


Figure 26: Percent missing plus not reached by item, block M1 (M2), Ireland, 2006 and 2009, position 4

Items are shown in the order they appear in the booklet

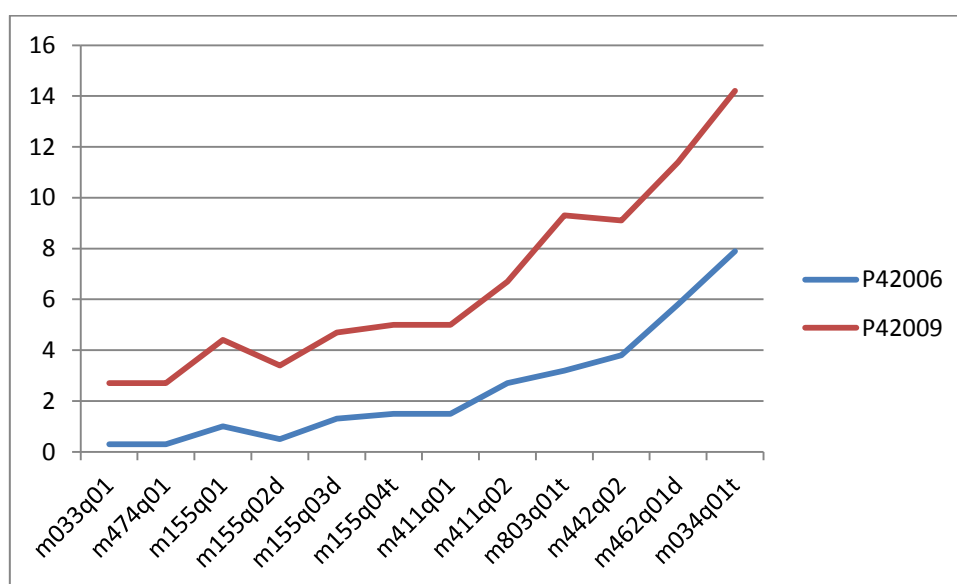


Figure 27: Percent not reached by item, block M1 (M2), Ireland, 2006 and 2009, position 4

Items are shown in the order they appear in the booklet

Analyses of Science Items, 2006 and 2009

Table 26 shows the percent of correct, missing and not reached responses for science blocks S1 (2009) and S4 (2006). It should be recalled that, unlike the previous analyses of mathematics and science, it was not possible to compare the same block across cycles; therefore comparisons are limited to block position *within* a cycle. The comparison countries as noted previously are Belgium, Hungary and the United Kingdom.

There is a 7-8% decline in percent correct for S4 in 2006 when one compares percent correct across countries for positions 1 and 4; and a decline of 7-10% for S1 in 2009. Ireland is not unusual in these

patterns of percent correct. Although comparing two different blocks of items, it can also be seen that percent missing and percent not reached in position 4 in particular tend to be quite stable in 2006 and 2009 and again, Ireland is no exception. Similarly, when one examines percent of incorrect responses in positions 1 and 4 within cycles, the pattern of results for Ireland (i.e. a slight increase in the percentage of incorrect attempted items) is not markedly different to that for the other countries considered.

Table 26: Average percent correct, missing and not reached for block S4 (2006) / block S1 (2009), positions 1 and 4, 2006 and 2009 – Belgium, Hungary, Ireland, and United Kingdom, and OECD averages

% Correct	P1 2006	P4 2006	P1 2009	P4 2009
Belgium	62.8	55.9	66.1	59.3
Hungary	61.1	54.4	65.5	55.2
Ireland	63.8	57.3	62.0	54.0
United Kingdom	61.8	53.7	61.7	53.8
OECD Average	59.8	50.9	64.3	53.1
% Incorrect				
Belgium	34.2	37.7	29.8	33.9
Hungary	36.1	38.6	30.3	36.2
Ireland	34.6	38.1	35.3	39.0
United Kingdom	36.0	39.3	35.2	39.2
OECD Average	37.0	41.5	31.7	39.3
% Missing				
Belgium	3.0	6.4	4.1	6.8
Hungary	2.8	7.0	4.2	8.6
Ireland	1.6	4.6	2.7	7.0
United Kingdom	2.2	7.0	3.1	7.0
OECD Average	3.2	7.6	4.0	7.6
% Not Reached				
Belgium	0.0	4.0	0.0	4.1
Hungary	0.0	3.5	0.0	1.7
Ireland	0.0	2.2	0.0	2.1
United Kingdom	0.2	2.9	0.2	2.3
OECD Average	0.0	5.2	0.1	5.7
% Missing + Not Reached				
Belgium	3.0	10.5	4.2	10.9
Hungary	2.8	10.4	4.2	10.3
Ireland	1.6	6.8	2.7	9.1
United Kingdom	2.4	9.9	3.3	9.2
OECD Average	3.2	12.8	4.1	13.3

Tables 27 and 28 show percent correct and missing/not reached for the countries shown in Table 26 by item type (note that the item classification in the case of science does not distinguish between short and long written response items). The overall patterns shown in Table 26 are mirrored in Tables 27 and 28. It is worth contrasting the findings for science with mathematics and particularly reading. That is, the marked declines in percent correct in Ireland (with a corresponding increase in missing responses) in position 4 are not evident for science. Unlike these other two domains, the response patterns for science of students in Ireland in 2006 and 2009 are not distinct to those from other countries or to the OECD averages.

Table 27: Average percent correct and missing and not reached for block S4 (2006) / block S1 (2009) multiple choice items, positions 1 and 4, 2006 and 2009 – Belgium, Hungary, Ireland, the United Kingdom, and OECD averages

% Correct	P1 2006	P4 2006	P1 2009	P4 2009
Belgium	61.5	55.7	71.7	64.3
Hungary	58.8	54.3	72.3	64.5
Ireland	62.2	56.7	67.9	60.9
United Kingdom	60.0	52.4	68.3	60.3
OECD Average	59.1	51.5	70.4	59.1
<hr/>				
% Incorrect				
Belgium	37.1	36.3	27.8	29.4
Hungary	40.1	38.1	27.4	32.7
Ireland	37.1	37.7	31.8	35.4
United Kingdom	38.7	40.3	31.1	36.3
OECD Average	39.6	38.9	29.0	33.4
<hr/>				
% Missing + Not Reached				
Belgium	1.4	8.0	0.5	6.3
Hungary	1.1	7.6	0.3	2.8
Ireland	0.7	5.6	0.3	3.7
United Kingdom	1.3	7.3	0.6	3.4
OECD Average	1.3	9.6	0.6	7.5

Since the blocks are not comparable across cycles, graphs comparing percent correct and missing across cycles and block position are not shown. However, Figure 28 shows the percent of not reached items in position 4 in 2006 and 2009. In contrast to reading and mathematics, there is no observable increase in the percentage of not reached items, which suggests that students engaged in the latter portion of the test when encountering a science block in a similar fashion in 2006 and 2009.

Table 28: Average percent correct and missing and not reached for block S4 (2006) / block S1 (2009) open response items, positions 1 and 4, 2006 and 2009 - Belgium, Hungary, Ireland, the United Kingdom, and OECD averages

% Correct	P1 2006	P4 2006	P1 2009	P4 2009
Belgium	67.2	56.3	58.2	52.2
Hungary	68.6	54.7	55.7	41.9
Ireland	68.8	59.2	53.5	44.2
United Kingdom	67.8	58.1	52.4	44.6
OECD Average	62.2	49.1	55.5	44.6
% Incorrect				
Belgium	24.5	25.2	32.4	30.2
Hungary	23.1	25.7	34.5	37.0
Ireland	26.8	29.9	40.4	39.1
United Kingdom	26.4	23.4	40.5	37.8
OECD Average	28.4	27.5	35.4	33.8
% Missing + Not Reached				
Belgium	8.3	18.5	9.4	17.6
Hungary	8.3	19.6	9.8	21.1
Ireland	4.4	10.9	6.1	16.7
United Kingdom	5.8	18.5	7.1	17.6
OECD Average	9.4	23.4	9.1	21.6

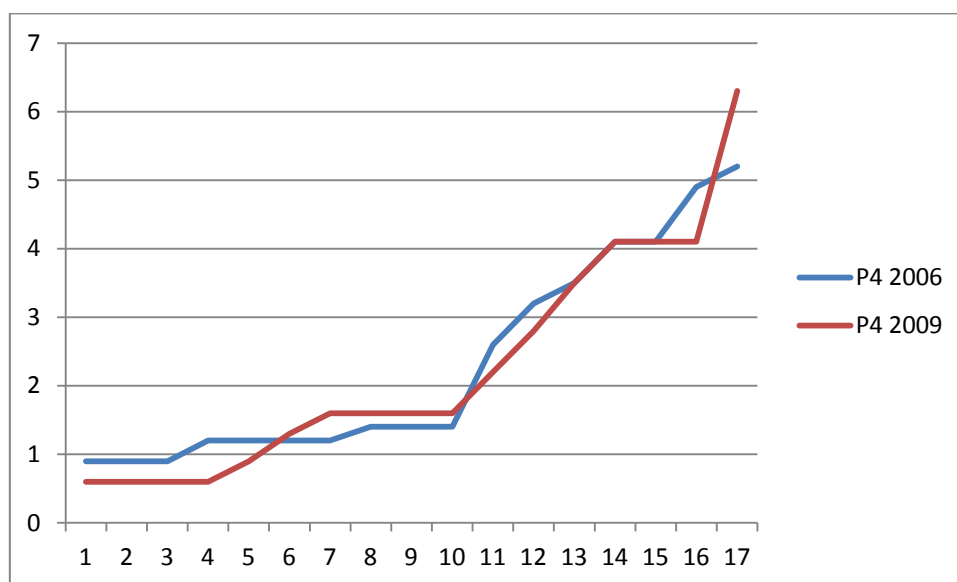


Figure 28: Percent not reached by item, block S4 (2006) / S1 (2009), Ireland, 2006 and 2009, position 4

Items are shown in the order they appear in the booklet

Summary and Conclusions

The analyses presented in this paper were conducted following recommendations in previous reports that documented and explored Ireland's decline in reading and mathematics, where an increase in the rate of not-attempted items and low student engagement in the PISA assessment in 2009 was observed in some testing sessions (Cosgrove et al., 2010; LaRoche & Cartwright, 2010).

Comparisons were made on the basis of two half-hour reading blocks that have been used to establish trends in reading in PISA 2003 and PISA 2009. Students' responses to reading items in PISA 2000 were not analysed since changes to the PISA test design between 2000 and subsequent cycles rendered comparisons problematic. In mathematics, students' responses on one of the mathematics blocks that has been used to establish trends using data from PISA 2006 and 2009 were analysed. Comparisons in the case of science were more limited due to the fact that the same items were not administered in intact blocks in 2006 and 2009, so two different blocks were selected for more general comparisons from these two years. For each of the three domains, three comparison countries were selected whose mean PISA scores did not differ to Ireland's in the reference cycle (2003 in the case of reading, and 2006 in the case of mathematics and science), and whose scores remained stable in PISA 2009 compared with the reference cycle. OECD averages were included as international benchmarks in this exercise.

The aim of the analysis was to examine and describe students' percentages of correct, incorrect, missing and not reached responses on common reading and mathematics blocks in 2003 and 2009 and on two different science blocks in 2006 and 2009. In the case of reading, more detailed analysis was conducted since this was the domain that has shown the greatest decline in achievement in the period considered, and the greatest decline in PISA scores since 2000 of all countries with data for both cycles (OECD, 2010e). Hence, for reading, response patterns by gender and socioeconomic status (SES) were analysed in addition to response patterns overall. Response patterns on reading items used for linking were also compared with reading items used for the first time in 2009 to see whether old and new items exhibited different patterns.

In all analyses, data are unweighted and significance tests were not conducted, so results and conclusions arising from them may be interpreted in terms of general patterns rather than as a confirmation or rejection of our hypothesis that the declines in achievement in reading and mathematics are due more to a decline in engagement on PISA than in the standards in these two domains.

A key to detecting changes in the response patterns of students lies in the PISA test design. Depending on the booklet assigned to a student, he or she may encounter a particular block of items in the first, second, third or fourth position of the test (the PISA test is two hours in duration, and each block is meant to require half an hour to complete, with a short break between the second and third blocks). It is generally found that responses to blocks of items appearing later in the test will have a lower percentage of correct responses and higher rates of missing responses due to what is termed 'student fatigue'. If student engagement was lower in 2009 relative to previous cycles, one would expect a more marked fatigue effect to be associated with responses to a block when it appears at the end of the PISA test (position 4) in 2009 compared to previous administrations of PISA. However, more generalised declines in proficiency (standards) would be evident in, say, a

reduction of percent correct and an increase in percent of incorrect and (to a lesser extent) missing responses on a block of items in all positions of the test.

The main findings and conclusions with respect to **reading** may be summarised as follows.

First, across the OECD on average, in Ireland, and the comparison countries, the percentage of correct, incorrect and missing responses has remained very stable for items administered in position 1 in 2003 and 2009. This is the case for both of the reading blocks examined (R1 and R2). This is indicative of stability in achievement levels, if we assume that performance on position 1 of the PISA test is representative of proficiency more generally.

Second, in position 4, there has been a decline in the percentage of correct responses across the OECD on average, with slight increases in missing and incorrect responses, between 2003 and 2009. This change is slightly more pronounced in the case of R1 than R2. The response patterns of students in Ireland in position 4 in 2003 and 2009 show the most marked changes in respect of the comparison countries (Belgium, the Netherlands and New Zealand) as well as the OECD averages. There has been a 10.8% decrease in the percentage of correct responses to R1 items in position 4, and a 13.4% decrease in the percentage of correct responses to R2 in this position between 2003 and 2009. This, however, is not accompanied by an increase in incorrect responses in the case of block R2; rather, there has been an increase in missing responses (from 6.1% to 10.2%) and not reached items (from 1.9% to 9.4%). For block R1, percent incorrect in position 4 has increased slightly from 23.9% to 28%, as have the percentages of missing responses (from 4.9% to 7.8%) and not reached responses (from 2.4% to 6.2%).

For block R2, it was shown that close to 6% of students did not attempt any items in position 4 in 2009 compared with 0% in 2000, and the percentages for block R1 are 1% and 4%, respectively. Thus, despite some variations in the response patterns observed in positions 1 and 4 in blocks R1 and R2 across cycles, it can be concluded that students in Ireland put considerably less effort into responding to reading items in position 4 than in position 1 in 2009 than in 2003, and that Ireland is unlike the comparison countries or the OECD averages in this respect. Indeed, Cartwright (2011) has shown that the pattern of missing and not reached responses in Ireland across PISA cycles represents an outlier among all other participating countries. That there has not generally been an increase in the percent of incorrect responses shows that the decrease in percent correct in position 4 is due to higher rates of missing/not reached items rather than higher rates of incorrect attempted items.

Together, these findings suggest that the marked decrease in percent correct in position 4 are more likely to be indicative of less effort or engagement than of lower levels of proficiency.

Third, students' response patterns in reading vary across cycles depending on item type in position 4 (but not in position 1). Although a decrease in the percentage of correct responses can be observed across all item types in position 4, the percentage of incorrect responses to written items is stable between 2003 and 2009, while it shows a slight increase in the case of multiple choice items. Why this is so cannot be inferred from the analyses. Also, comparisons by item type were made on the basis of block R2 only, so these changes may be related to items specific to this block than to multiple choice reading items more generally. However, given that incorrect responses to written items have remained stable in position 4, this does suggest that the increase in incorrect responses

to multiple choice items may be due to more guessing on the part of students in 2009 relative to 2003, as opposed to a decline in ability to respond correctly to questions.

Fourth, that the decline in student percent correct for R2 (coupled with the large increase in missing responses) is more marked than for R1 suggests that while student disengagement may be a general issue in PISA 2009, for some reason, this appears to have been more problematic for block R2. Again, why this is the case cannot be inferred from the analyses presented here, but it may be the case that the specific content of block R2 is now considerably less interesting or engaging to students than it was in 2003, relative to block R1. It is also important to note that in 2009, R2 in position 4 was preceded by a science block, while R1 was preceded by another reading block, and this may have had a differential impact on students' engagement with the reading items that they encountered in the last position.

Fifth, the comparison of positioning effects for one of the link clusters (R2) and one of the four new reading clusters (R3A) indicated that positioning effects were stronger for R2 than for R3A. For example, percent correct declined by almost 18% between positions 1 and 4 in the case of R2, but by only 8% in the case of R3A. This suggests that positioning effects may be stronger for the link items compared with the new reading items that were used in PISA 2009, but an analysis of all reading items administered in PISA 2009 would be necessary to conclude this more definitively.

Sixth, changes in the percent correct, incorrect, missing and not reached responses varied depending on the SES and gender of respondents. Stronger positioning effects appeared to be associated with low-SES students and boys when one compares responses to R2 in 2003 and 2009. Related to this, there has been a more marked increase in the percentage of other language speakers among boys than girls between 2003 and 2009, although the analyses do not address whether or not these shifts in the compositions of sub-groups are related to response patterns on the PISA assessment of reading.

The analyses of responses to the **mathematics** indicated the following key findings and conclusions.

First, percent correct on mathematics items across the OECD on average has remained stable across 2006 and 2009 in both positions 1 and 4, though as expected, is lower in both cycles, by about 5-6%, in position 4. In Ireland, percent correct in mathematics has decreased by 3.2 percentage points in position 1 and by 4.7 percentage points in position 4. The percentages of incorrect responses have remained relatively stable across 2006 and 2009 in Ireland, the three comparison countries (Iceland, Poland and Slovenia), and across the OECD on average. However, Ireland shows an increase in missing responses that is larger than the comparison countries or the OECD averages in both position 1 (from 6.9% to 8.7%) and position 4 (from 8.9% to 12.0%). Similarly, the percentages of not reached items in position 4 are generally quite stable across 2006 and 2009 in the comparison countries and across the OECD on average, but the rate for Ireland has increased from 2.5% to 6.6%.

Thus, this first set of findings provides some evidence for *both* a decline in proficiency (for example, with a decrease in the percentage of correct responses in position 1) *and* an increase in disengagement (for example, with an increase in position 4 of missing and not reached responses combined from 11.4% to 18.6%).

Second, response patterns across cycles varied depending on item type. In position 1, for example, percent correct in Ireland declined by 2.1% on multiple choice items, 6.3% on short response items, and did not change at all on open response items (i.e. those requiring a longer written response). Since comparisons of item types are limited to just one mathematics block, it is not possible to generalise these findings to all PISA mathematics items; however, it is possible that the differential changes in percentages of correct responses may be more related to the mathematics content assessed by items as opposed to the item type *per se*.

Third, relative to reading, response patterns in mathematics in all positions are relatively more stable. Nonetheless, however, there is a marked increase between 2006 and 2009 in the percentages of missing and not reached responses in position 4.

Analyses of **science** items were more limited due to the fact that the same intact block of items was not administered in 2006 and 2009. However, the results of analyses of percent correct, incorrect, missing and not reached responses in positions 1 and 4 in 2006 and 2009 indicated first that science generally appears to be less prone to positioning effects than reading or mathematics. Second, in contrast to reading and mathematics, Ireland's response patterns were not markedly different to the three comparison countries (Belgium, Hungary and the United Kingdom) or on average across the OECD in either 2006 or 2009, in either positions 1 or 4. So the main conclusion that may be drawn with respect to science is that the stable response patterns for Ireland are consistent with the published results which show no change in the science scores of students in Ireland in between 2006 and 2009.

This, however, begs the question as to why response patterns in science have remained so stable in Ireland when response patterns on common items in mathematics and even more so reading have fluctuated so markedly, particularly in position 4. The present paper cannot address this question, but several possibilities arise. First, science blocks tend to contain more items than reading or mathematics and this may render response patterns inherently more stable in the case of science. Second, proportionately fewer science items require a written response. For example, in 2009, mathematics and reading contained more written response items (54% and 52% respectively) than science (24%). The analyses presented here showed that students' response patterns tended to be more stable on multiple choice items than on written response items. Third, the nature of the written responses required of PISA science items tended to be short, compared with some reading and mathematics items, which require, by comparison, an extensive written response. Cosgrove and Moran (2011) analysed response patterns on the PISA 2009 print and digital assessments and, consistent with Cartwright (2011), they conclude that differences between domains in terms of the distribution of response types and different levels of effort required to complete these may be acting as aspects of the PISA assessment design that are confounded with the measurement of students' achievement as intended, i.e. a reliable indicator of underlying proficiency.

Finally, it was noted elsewhere (Cosgrove et al., 2011; LaRoche & Cartwright, 2011; Perkins et al., 2011) that for the first time in any PISA administration, eight very low-performing schools participated in PISA 2009. The contributions of these 'outlier' schools to response patterns has not been examined in the present paper, and readers are referred to Cosgrove and Moran (2011) and Perkins et al. (2012) for further analyses of these schools.

Unfortunately, neither this set of analyses nor any other will be able to address the question as to what extent the declines in PISA 2009 reading and mathematics represent 'real' declines in student proficiency. This is explained in Cartwright (2011), where it is noted that aspects relevant to comparing trends in achievement interact with one another, and that measurement and link errors further compound the issue. However, with successive cycles of PISA, as more information on achievement becomes available, we can begin to establish a more firm and reliable basis on which to draw conclusions about trends in students' achievements on PISA over time.

References

- Cartwright, F. (2011). *PISA in Ireland, 2000-2009: Factors affecting inferences about changes in student proficiency over time*. Dublin: Educational Research Centre.
- Cosgrove, J., & Moran, G. (2011). *Taking the PISA 2009 test in Ireland: Students' response patterns on the print and digital assessments*. Dublin: Educational Research Centre.
- Cosgrove, J., Perkins, R., Moran, G., & Shiel, G. (2011). *Digital reading literacy in the OECD Programme for International Student Assessment (PISA 2009): Summary of results for Ireland*. Dublin: Educational Research Centre.
- Cosgrove, J., Shiel, G., Archer, P., & Perkins, R. (2010). *Comparisons of performance in Ireland on PISA 2000 and PISA 2009: A preliminary report to the Department of Education and Skills*. Dublin: Educational Research Centre.
- LaRoche, S., & Cartwright, F. (2010). *Independent review of the 2009 PISA results for Ireland: Report prepared for the Educational Research Centre at the request of the Department of Education and Skills*. Dublin: Department of Education and Skills.
- OECD (2009). *PISA data analysis manual (2nd edition)*. Paris: Author.
- OECD (2010a). *PISA 2009 results: What students know and can do – Student performance in reading, mathematics and science (Volume I)*. Paris: Author.
- OECD (2010b). *PISA 2009 results: Overcoming social background – Equity in learning opportunities and outcomes (Volume II)*. Paris: Author.
- OECD (2010c). *PISA 2009 results: Learning to learn – Student engagement, strategies and practices (Volume III)*. Paris: Author.
- OECD (2010d). *PISA 2009 results: Resources, policies and practices (Volume IV)*. Paris: Author.
- OECD (2010e). *PISA 2009 results: Learning trends – Changes in student performance since 2000 (Volume V)*. Paris: Author.
- OECD (2011). *PISA 2009 results: Students on line – Digital technologies and performance (Volume VI)*. Paris: Author.
- Perkins, R., Moran, G., Cosgrove, J., & Shiel, G. (2010). *PISA 2009: The performance and progress of 15-year-olds in Ireland – Summary Report*. Dublin: Educational Research Centre.
- Perkins, R., Moran, G., Shiel, G., & Cosgrove, J. (2011). *Reading literacy in PISA 2009: A guide for teachers*. Dublin: Educational Research Centre.
- Perkins, R., Cosgrove, J., Moran, G., & Shiel, G. (2012). *PISA 2009: Results for Ireland and changes since 2000*. Dublin: Educational Research Centre.
- Shiel, G., Moran, G., Cosgrove, J., & Perkins, R. (2010). *A summary of the performance of students in Ireland on the PISA 2009 test of mathematical literacy and a comparison with performance in 2003: Report requested by the Department of Education and Skills*. Dublin: Educational Research Centre.