Taking the PISA 2009 Test in Ireland:

Students' Response Patterns on the Print and Digital Assessments

Jude Cosgrove

Gráinne Moran

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

December 2011

Taking the PISA 2009 Test in Ireland:

Students' Response Patterns on the Print and Digital Assessments

Jude Cosgrove

Gráinne Moran

Table of Contents

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 the subject of a presentation to the Joint Oireachtas Committee on Education and Skills as well as a debate of the Joint Committee in May 2011[1].

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 are 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[2]. This was evidenced in their behaviour during some of the observed testing sessions 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 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 two which provides a more in-depth look at students' response patterns on the PISA tests. It examines students' response patterns on the PISA digital and print assessments in 2009. The other (Cosgrove, 2011) examines students' response patterns on the print assessments in PISA 2003, 2006 and 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 (2011) and Cartwright (2011) give a comprehensive picture to students' response patterns on the PISA assessments. They are summarised in Chapter 9 of the main 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 these aspects of PISA are referred to OECD (2010a, b, c, d, e; 2011), as well as to Perkins et al. (2012). Readers who wish to become more familiar with the technical aspects of PISA are referred to the *PISA 2009 Technical Report* (OECD, 2011). Details of PISA publications can be found at www.erc.ie/pisa and www.pisa.oecd.org.

[1] http://debates.oireachtas.ie/EDJ/2011/01/13/00004.asp

[2] In this report, 'engagement' is intended in a general sense, i.e. students' willingness to engage with the PISA assessment, to attempt both easy and challenging items, and to persist in their efforts to the end of the test.

Aim of the Present Study

To describe students' responses and response patterns by item type, taking block position into account, and drawing on both paper-based and digital item data. The study will be focussed on reading, though mathematics and science are also examined.

Research Questions

1. What is the percent correct and percent missing on the PISA 2009 assessment, by domain, mode (print versus digital) and item type (multiple choice versus written response)?
2. How do percent correct and percent missing by domain, mode and item type vary by student gender, grade level, economic, social and cultural status (ESCS), school sector, and school SSP status[3]?
3. What percentage of variation in percentages correct and missing is accounted for by the student and school characteristics in (2), and does this vary by domain, mode or item type?
4. Are there additional variations evident depending on the position of items in the test form, and does this vary, depending on mode (print versus digital) and item type?

Method

Participants

For the first part of the analysis, which examines students' response patterns across domains and item types, we include the responses of all students, where available. While all students attempted some print reading items (N=3937), not all students attempted mathematics, science and digital reading items. Specifically, data are available for 2749 students for mathematics 69.8%, for 2745 students for science (69.7%), and for 1409 students for digital reading (35.8%). Data are available for 669 students (17.0%) in all four domains.

For the second part of the analysis, i.e. that which examines positioning effect, we are restricted to students who attempted a specific reading cluster (R3A, the selection of which is explained more in the following section), which appears in 4 of the 13 print assessment booklets (Booklets 1, 3, 4 and 6). In all, 1176 students (29.8%) of students attempted one of these four booklets, and the digital reading assessment was attempted by 425 of them (35.5%).

Selection of Booklets

Because the PISA test design entails blocks rotated across booklets, it was decided to sample a subset of four of the 13 test booklets for the purpose of analysing position effects. Block R3A was selected as the focus of the analysis, since three of the four units in this block have been publicly released, thereby allowing illustrations of sample questions and student responses where appropriate (see Perkins et al., 2012, Appendix B, for the released print reading items from PISA 2009). Block R3A also has a good mix of item types. Table 1 shows the test design for PISA 2009. The reading block selected for detailed analysis is highlighted in the table (Booklets 1, 3, 4 and 6).

[3] 'SSP status' refers to the Department of Education and Skills' School Support Programme, i.e. whether or not a school is in receipt of additional support to address educational disadvantage under the Delivering Equality In Schools (DEIS) initiative (Department of Education and Science, 2005).

Note that the PISA booklet design does not conclusively allow us to separate positioning effects from preceding domain effects, since both change at the same time.[4] For example, if block R3A were to appear in position 4 three times, preceded by a reading, mathematics and science block in position 3, then some disentanglement of this issue would be possible, but the design is such that each block appears in each position just once. It should also be noted that each print block is designed to require approximately half an hour to complete, with a total testing time of two hours, excluding a short break after the first hour (first two blocks). However, in practice, students are free to progress to the second half of their test booklets before the first hour of testing time is finished, should they wish to do so.

Table 1: Test design for the PISA 2009 print assessment

Booklet	Cluster			
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

The test design for the digital reading assessment is shown in Table 2. Unlike the print test design, only one domain is involved, and as such it is a more balanced design, with each of three reading blocks appearing in each position, and each block paired with each of the other two blocks. Also, the testing time for the digital assessment was shorter than for the print assessment, with a total testing time of 40 minutes (or 20 minutes per block, with no break between blocks) and again in theory, students could have progressed to the second block of the digital reading test before the first 20 minutes' of testing time had elapsed.

[4] By 'positioning effect', we mean the tendency for responses to a block of items that appears later in a booklet to have a lower percent correct compared with appearing earlier in a booklet due to student characteristics such as fatigue or decreased engagement. The magnitude of positioning effects can vary depending on factors such as test length, item difficulty, and whether the test is low-stakes or not. By 'domain effect', we mean an hypothesised association between the responses to a specific block (in this case, reading) and the content and perceived difficulty of the preceding block; for example, an easy preceding science block may promote student engagement with the subsequent reading block, whereas as difficult preceding mathematics block may serve to disengage or discourage students from engaging with the subsequent reading block.

Table 2: Test design for the PISA 2009 digital reading assessment

Form ID	Cluster	
1	A	B
2	B	A
3	B	C
4	C	B
5	C	A
6	A	C

Analysis

For the initial descriptive analyses that do not take positions of items into account, all students were included and results generated using replicate weights, with standard errors corrected for sampling and measurement error (OECD, 2009). For the analyses of the subset of students that attempted Block R3A, the replicate weights have not been applied, since they were computed on the basis of the whole sample; however, student sample weights have been used. Therefore, analyses that examine position effect have standard errors that may be somewhat underestimated since sampling variation has not been fully taken into account.

Following descriptive analyses which compare percent correct and missing of students by various sub-groups such as gender and grade level, we present the results of multiple regression analyses. The specific details of these analyses are provided immediately prior to presenting the results. The purpose of the regression analyses is to bring together the descriptive analyses in order to provide a synthesis of students' response patterns.

In analyses of students' responses to test items, it is common to distinguish between two types of missingness. That is, an item is 'missing' if a student did not answer it but *did* respond to one or more questions later in the test booklet, while it is 'not reached' if a student did not answer it and *did not* respond to any questions later in the test booklet. In the present analyses, however, we have combined missing and not reached responses into a single 'missing' category, mainly due to the fact that the percentage of not reached items tends to be very small (i.e. most students had adequate time to complete the test).

Also in the present analyses, percent correct differentiates between partially and fully correct responses in the case of partial credit items (i.e. is based on total score points), while percent missing is computed on the basis of the total number of items attempted by students, rather than the total number of score points.[5]

[5] As an example, take a test booklet with 20 questions, four of which are partial credit and can be scored 0, 1 or 2. The total number of score points is thus 24, while the number of items is 20.

Results

Characteristics of Test Items

For the purposes of the present analyses, we distinguish between only two item types – those requiring a written response and those that are multiple choice in nature. Table 3 shows the international item percentages correct and item difficulties[6] overall and by item type by block and domain.

The table shows that there is more variability across clusters in terms of item difficulty in the case of print reading relative to science and particularly mathematics. In fact, the cluster selected for detailed analysis (R3A) is the most difficult cluster of the seven reading clusters (with an average percent correct of 52.8%, compared to an average of 60.2% across all reading items).

Table 3. International item percent correct and difficulty (delta) by item format, cluster, domain, and overall, PISA 2009

Cluster/Domain	Multiple Choice					Written Response					All Item Types				
	N items	% Correct		Delta		N items	% Correct		Delta		N items	% Correct		Delta	
		Mean	SD	Mean	SD		Mean	SD	Mean	SD		Mean	SD	Mean	SD
Print Reading															
R1	6	73.4	12.6	-0.941	0.857	6	53.7	18.4	0.231	1.006	12	63.5	18.2	-0.355	1.081
R2	4	64.4	12.5	-0.393	0.751	10	52.2	19.6	0.353	1.146	14	55.7	18.3	0.140	1.078
R3A	10	56.9	12.6	0.223	0.682	5	44.6	17.8	0.971	1.104	15	52.8	15.1	0.472	0.883
R4A	8	61.0	15.1	-0.025	0.894	8	62.4	20.8	-0.125	1.164	16	61.7	17.6	-0.075	1.004
R5	8	53.9	18.8	0.345	1.060	7	52.5	19.5	0.380	1.053	15	53.2	18.4	0.361	1.019
R6	7	61.4	16.7	-0.176	0.976	8	58.0	22.6	0.070	1.305	15	59.6	19.4	-0.045	1.130
R7	5	54.6	37.5	0.158	2.680	9	76.6	13.4	-1.013	0.928	14	68.8	25.8	-0.595	1.755
Mathematics															
M1	3	58.3	15.5	-0.639	0.921	9	43.0	21.5	0.180	1.154	12	46.8	20.7	-0.025	1.122
M2	6	60.5	16.9	-0.829	1.049	6	36.0	22.6	0.742	1.461	12	48.3	23.0	-0.043	1.464
M3	7	50.6	13.0	-0.246	0.731	4	36.0	21.1	0.635	1.245	11	45.3	17.0	0.074	0.992
Science															
S1	10	64.6	13.4	-0.582	0.769	7	49.0	17.5	0.250	1.008	17	58.2	16.7	-0.240	0.944
S2	13	54.9	17.8	-0.068	1.015	5	39.1	12.5	0.708	0.610	18	50.5	17.7	0.148	0.971
S3	12	54.4	17.4	-0.020	0.964	6	50.0	14.5	0.277	0.691	18	52.9	16.2	0.079	0.873
Digital Reading															
A	9	56.9	19.3	0.231	0.970	1	47.3	0.0	0.689	0.000	10	55.9	18.4	0.277	0.926
B	7	65.2	14.7	-0.280	0.814	3	51.0	42.4	1.838	1.035	10	58.4	17.5	0.115	0.977
C	6	76.0	15.5	-0.997	1.184	3	50.3	48.5	0.802	0.688	9	66.8	18.4	-0.435	1.260
All Mathematics	16	55.8	14.8	-0.538	0.877	19	39.3	20.8	0.453	1.231	35	46.8	19.9	0.000	1.180
All Print Reading	48	60.2	18.3	-0.060	1.176	53	58.2	20.4	0.054	1.187	101	59.2	19.4	0.000	1.177
All Science	35	57.5	16.7	-0.198	0.939	18	46.6	15.2	0.386	0.793	53	53.8	16.9	0.000	0.928
All Digital Reading	22	64.8	18.0	-0.266	1.068	7	45.7	8.5	1.838	0.837	29	60.2	18.1	0.000	1.062

[6] In the case of item deltas, these have been standardised to have a mean of zero within domains, and hence are not useful for comparing difficulty across domains (but are fine for comparing difficulties within a domain).

In the case of mathematics and science, written response item types tend to be more difficult than multiple choice items. This is not so in the case of reading, where, on average across all clusters, multiple choice questions (60.2% correct) are similar in difficulty to written response questions (58.2% correct). However, this pattern varies across the reading clusters. For example, written response items in cluster R7 (76.6% correct) are much easier than multiple choice questions in this cluster (54.6% correct), while the opposite is the case in clusters R1, R2 and R3A. The digital reading items requiring a written response are, on average, more difficult (45.7% correct) than the multiple choice items for this domain (64.8% correct), though average item difficulty by item format varies across the three digital reading blocks.

It can also be seen that while approximately half of the reading and mathematics items are multiple choice, with the remainder requiring a written response, fewer of the science items require a written response (i.e. 34%, or 35/53 items). The digital reading assessment also has fewer written response items, at about 24% (7/29 items).

Percent Correct and Percent Missing – Descriptive Analyses

Table 4 shows, for all students in Ireland who took part in PISA 2009, percent correct and missing for all items, and by item type, for print reading, mathematics, science, and digital reading.

Overall student percent correct is highest for print and digital reading (58.7% and 58.4%, respectively), then science (54.1%), and is lowest for mathematics (43.6%). Students also missed more of the mathematics items overall, with an average of 11.8% missing on this domain. Missing responses are second highest for print reading (8.5%), while the percent missing for science and digital reading are lower (6.3% and 6.5%, respectively).

Table 4: Student percent correct and percent missing, overall and by item type, for print reading, mathematics, and digital reading, PISA 2009 (Ireland)

Domain	All item types				Written Response				Multiple Choice			
	Correct		Missing		Correct		Missing		Correct		Missing	
	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
Print Reading	58.7	0.61	8.5	0.48	58.3	0.65	12.6	0.61	58.3	0.61	4.3	0.40
Mathematics	43.6	0.56	11.8	0.45	35.0	0.63	19.4	0.61	54.3	0.60	3.8	0.33
Science	54.1	0.61	6.3	0.47	49.5	0.71	10.6	0.63	57.1	0.61	4.1	0.42
Digital Reading	58.4	0.71	6.5	0.36	43.1	0.83	12.9	0.63	67.7	0.73	4.1	0.31

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

Percent correct on written responses is also lowest for mathematics (35.0%), while it is highest for print reading (58.3%), and percent correct for science and digital reading are intermediate to mathematics and print reading (49.5% and 43.1%, respectively). Percent correct for all three print domains on multiple choice items are similar, ranging from 54.3% (mathematics) to 58.3% (reading), while percent correct on digital reading multiple choice items is higher (67.7%).

In all domains, for multiple choice items, percent missing is low, at around 4%. However, there is more variability in the percent of missing written responses across domains, and this may reflect a combination of factors, including the domain itself, average item difficulty of written responses in each domain, and the overall proportion of written response items in each domain, amongst other

things. In all domains, however, percent of missing responses is higher on written response items compared with multiple choice ones, ranging from 10.6% for science to 19.4% for mathematics, with print and digital reading intermediate to science and mathematics, at around 13%.

The next part of this section takes the same data shown in Table 4 and compares the results for subgroups of interest, in order to identify diverging patterns in percent correct and percent missing among these subgroups that can then inform the regression analyses that follow.

Table 5 shows student percent correct and missing by domain and item type for gender and school type combined (Tables A1 and A2 show these estimates separately for gender and school type). Variations across both gender and school type can be observed. Taking print reading first, the student average percent correct for boys ranges from a low of 51.2% in vocational schools to highs of 58.3% mixed secondary schools and 57.9% in single sex secondary schools. For girls, this ranges from 54.4% (vocational schools) to 65.1% (single sex secondary schools). For all reading items, the pattern of missing responses is consistent with percent correct scores, i.e. girls tended to miss fewer items than boys, though 10% of items were not attempted by girls in vocational schools. Percent correct does not vary appreciably across item type for reading, when comparing across gender/school type. Although percent of non-attempted reading items are higher for questions requiring a written response across all groups, patterns are again consistent with the percentages of missing items for all items combined.

Turning now to mathematics, it can be seen that boys have a higher percent correct across mathematics items than girls, although the size of the gender difference depends on the sector. The biggest disparity is associated with vocational schools (with 36.6% correct for girls and 41.1% correct for boys) and the smallest is associated with single sex secondary schools (with 44.5% correct for girls and 45.4% correct for boys). The percentage of missing responses on mathematics items varies across gender/school type, with the highest rates of non-attempted questions by both boys and girls in vocational schools, and lowest rates by girls in mixed and single-sex secondary schools.

When comparing percent correct for mathematics across the two item types, some interesting results emerge. For example, girls in community/comprehensive and vocational schools do better on written items than might be expected, given their overall percentages correct. And, consistent with reading, the percentage of missing responses is higher on written response than on multiple choice items, with the rate of missing written responses ranging from around 15% for girls in single sex and mixed sex secondary schools, to 26.2% for boys in vocational schools.

In the case of science, the percentages of correct responses across all items do not vary as much by gender/school type as they do for print reading or mathematics. The percent correct for both boys and girls in vocational schools (49.1% and 47.2%, respectively) is the lowest across all groups. The percentage of missing science items tends to be lower for all groups, compared with print reading and mathematics, and again is higher for written response than multiple choice items across all groups in Table 5.

Comparisons of percent correct on print and digital reading indicate that both boys and girls in community/comprehensive schools did better on the digital reading items, while students in mixed secondary schools did better on print reading items. In vocational schools, boys did slightly better on print reading questions, while girls did better on digital reading items. There are little differences

between print and digital reading (percent correct) for boys and girls in single-sex secondary schools. Across all school sectors and for both genders, the percentage of missing items was lower for digital reading than for print reading. This difference is more marked in community/comprehensive and vocational schools.

Table 5: Student percent correct and percent missing, overall and by item type, for print reading, mathematics, and digital reading, PISA 2009 – comparisons by student gender and school type (Ireland)

	All Items				Written Items				Multiple Choice Items			
	Correct		Missing		Correct		Missing		Correct		Missing	
Print Reading	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
Community/ Comprehensive: Boys	52.7	1.26	12.6	1.49	51.7	1.54	18.1	1.80	53.2	1.11	7.0	1.34
Community/ Comprehensive: Girls	60.5	2.05	6.3	1.06	60.4	2.31	9.2	1.46	59.7	1.75	3.4	0.74
Vocational: Boys	51.2	1.52	13.6	1.56	50.7	1.58	20.1	1.81	50.9	1.55	6.9	1.46
Vocational: Girls	54.4	1.65	10.0	1.43	54.0	1.67	15.0	1.80	54.0	1.79	5.2	1.38
Boys' Secondary	57.9	1.88	9.1	1.32	57.5	1.92	13.6	1.73	57.5	1.87	4.7	0.97
Girls' Secondary	65.1	0.89	4.6	0.37	65.3	0.92	7.0	0.55	64.2	0.96	2.2	0.27
Mixed Secondary: Boys	58.3	1.55	9.2	1.17	57.2	1.88	13.8	1.68	58.9	1.41	4.4	0.81
Mixed Secondary: Girls	63.6	1.21	5.2	0.68	63.7	1.14	7.8	0.88	62.6	1.34	2.7	0.58
Mathematics												
Community/ Comprehensive: Boys	44.2	1.68	12.7	0.96	34.0	1.79	20.8	1.22	56.5	2.41	4.0	1.19
Community/ Comprehensive: Girls	42.1	1.72	11.8	1.31	35.1	2.31	19.3	1.77	51.9	1.70	3.4	0.99
Vocational: Boys	41.1	1.54	17.0	1.24	31.6	1.44	26.2	1.60	52.2	1.80	7.0	1.08
Vocational: Girls	36.6	1.39	13.9	1.67	29.3	1.57	22.7	2.30	46.2	1.84	4.4	1.19
Boys' Secondary	45.4	1.29	12.0	0.97	36.4	1.39	20.0	1.48	57.2	1.56	3.7	0.61
Girls' Secondary	44.5	1.04	8.8	0.55	36.2	1.14	15.0	0.90	54.6	1.13	2.4	0.37
Mixed Secondary: Boys	47.7	1.21	11.7	1.23	39.0	1.57	19.2	1.93	58.2	1.35	4.0	0.82
Mixed Secondary: Girls	44.1	1.50	8.4	1.18	36.0	1.91	14.7	1.89	54.3	1.76	1.7	0.41
Science												
Community/ Comprehensive: Boys	53.2	1.73	8.7	2.06	49.3	1.99	12.5	2.42	55.7	1.68	6.8	1.89
Community/ Comprehensive: Girls	52.7	2.23	5.4	1.17	47.4	2.84	10.4	1.76	56.2	2.17	3.0	0.93
Vocational: Boys	49.1	1.67	10.3	1.84	45.6	1.91	16.5	2.12	51.4	1.79	7.2	1.80
Vocational: Girls	47.2	1.52	8.3	1.42	40.9	1.92	14.5	2.16	51.0	1.49	5.1	1.14
Boys' Secondary	55.7	1.60	6.3	0.84	50.5	1.83	11.1	1.29	58.9	1.60	3.9	0.67
Girls' Secondary	57.5	1.09	3.6	0.41	53.1	1.35	6.3	0.74	60.3	1.01	2.2	0.29
Mixed Secondary: Boys	57.0	1.91	6.3	0.93	52.8	2.33	10.1	1.32	59.6	1.75	4.3	0.77
Mixed Secondary: Girls	56.6	1.48	3.6	0.57	51.7	1.90	6.6	0.96	59.5	1.41	2.2	0.46
Digital Reading												
Community/ Comprehensive: Boys	56.2	1.71	8.5	0.78	41.0	2.13	16.8	1.49	65.1	1.70	5.5	1.12
Community/ Comprehensive: Girls	63.5	2.39	4.3	0.65	48.0	2.89	9.9	1.63	72.5	2.35	2.2	0.47
Vocational: Boys	49.7	1.29	10.6	1.34	34.2	1.60	19.8	2.00	59.3	1.48	7.1	1.24
Vocational: Girls	56.3	2.17	6.1	0.73	41.2	2.47	12.5	1.54	65.3	2.16	3.7	0.67
Boys' Secondary	57.4	2.13	7.2	0.95	41.6	2.42	14.7	1.83	67.2	2.25	4.4	0.69
Girls' Secondary	64.6	1.12	4.0	0.44	49.8	1.63	7.7	0.81	73.3	0.96	2.6	0.37
Mixed Secondary: Boys	56.0	1.90	7.4	0.83	39.5	2.03	15.1	1.90	65.8	1.95	4.5	0.73
Mixed Secondary: Girls	62.0	2.25	4.4	0.86	47.5	3.03	7.6	1.18	70.7	2.02	3.2	0.88

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

Table 6 shows item statistics for students across grade levels[7]. Across all domains, there is a similar pattern of overall percent correct responses, i.e. increasing from Grade 8 to Grade 10, then dipping slightly at Grade 11. The overall percentages of missing responses mirror the pattern associated with percent correct across all domains. This pattern holds across domains for both written and multiple choice items, although the differences across grade levels are more marked in the case of written response items. The particularly high rate of missing written responses in all domains by Grade 8 students is noteworthy. It should be noted that just 2.4% of students that participated in PISA 2009 were in Grade 8.

Table 6: Student percent correct and percent missing, overall and by item type, for print reading, mathematics, and digital reading, PISA 2009 – comparisons by grade level (Ireland)

	All Items				Written Items				Multiple Choice Items			
	Correct		Missing		Correct		Missing		Correct		Missing	
Print Reading	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
Grade 8	33.4	2.62	23.9	2.76	32.8	3.22	32.7	3.87	34.3	2.32	14.5	2.57
Grade 9	57.1	0.71	8.9	0.56	56.5	0.81	13.5	0.74	56.9	0.67	4.3	0.45
Grade 10	64.6	1.00	6.1	0.48	64.5	1.09	9.2	0.67	64.0	0.95	3.2	0.39
Grade 11	59.7	1.24	8.1	1.01	59.9	1.31	11.5	1.24	58.6	1.25	4.6	0.86
Mathematics												
Grade 8	25.0	2.83	28.2	2.95	17.5	3.22	41.7	4.01	33.8	3.73	12.8	2.78
Grade 9	42.0	0.74	12.3	0.57	33.5	0.79	20.4	0.82	52.7	0.84	3.6	0.37
Grade 10	47.7	0.95	9.4	0.60	39.1	1.21	15.5	0.99	58.3	1.08	3.3	0.52
Grade 11	45.9	1.12	11.4	0.91	37.0	1.24	18.3	1.41	57.2	1.39	4.0	0.68
Science												
Grade 8	33.2	2.61	18.9	2.89	25.9	3.49	30.8	3.68	37.8	2.38	13.2	2.96
Grade 9	53.2	0.75	6.4	0.59	48.2	0.89	11.0	0.83	56.3	0.76	4.2	0.51
Grade 10	58.4	1.00	4.2	0.41	54.3	1.24	7.5	0.67	61.0	0.99	2.5	0.35
Grade 11	55.3	1.24	6.5	0.98	51.6	1.49	10.1	1.39	57.9	1.28	4.6	0.85
Digital Reading												
Grade 8	35.1	3.67	16.8	2.47	20.5	3.97	33.7	4.48	43.8	3.70	10.3	2.51
Grade 9	56.2	0.83	7.2	0.47	40.7	0.96	13.5	0.78	65.5	0.89	4.8	0.43
Grade 10	65.4	1.58	3.8	0.56	50.6	1.91	8.8	1.06	74.3	1.49	1.9	0.44
Grade 11	61.0	1.40	6.1	0.72	45.0	1.57	12.5	1.40	70.6	1.52	3.6	0.66

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

[7] Grade 8 = Second Year, Grade 9 = Third Year, Grade 10 = Transition Year, Grade 11 = Fifth Year.

Table 7 shows the percent correct and missing for all domains, overall and by item type, for students with high, medium and low scores on the ESCS scale.[8] There is a fairly monotonic increase in percent correct across all domains as ESCS increases, and this is mirrored by a monotonic decrease in the percentage of missing responses. This is true for both multiple choice and written responses, though again, missing response rates are higher for written response items than for multiple choice ones, regardless of the domain. It may also be noted that the difference in average percent correct on print reading between the high and low ESCS groups (16.2%) is larger than the difference on digital reading (13.1%), indicating a weaker association between ESCS and percent correct on digital reading than on print reading.

Table 7: Student percent correct and percent missing, overall and by item type, for print reading, mathematics, and digital reading, PISA 2009 – comparisons ESCS group (Ireland)

	All Items				Written Items				Multiple Choice Items			
	Correct		Missing		Correct		Missing		Correct		Missing	
Print Reading	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
Low	51.0	0.79	11.4	0.67	50.3	0.93	17.4	0.87	51.0	0.74	5.6	0.63
Medium	59.1	0.86	7.9	0.66	59.0	0.90	11.8	0.79	58.1	0.87	4.0	0.59
High	67.2	0.66	5.0	0.38	67.0	0.77	7.4	0.53	66.7	0.65	2.6	0.29
Mathematics												
Low	35.5	0.71	16.3	0.72	26.1	0.83	27.1	1.06	46.8	0.89	5.2	0.57
Medium	44.7	0.78	11.0	0.57	36.4	0.87	18.5	0.86	55.0	0.95	3.1	0.43
High	50.8	0.77	7.7	0.47	42.8	0.91	12.3	0.69	61.4	0.94	2.5	0.37
Science												
Low	46.8	0.82	8.5	0.69	42.1	0.88	14.5	0.92	49.8	0.91	5.4	0.65
Medium	54.9	0.88	5.4	0.50	50.4	1.07	9.3	0.77	57.7	0.86	3.4	0.43
High	62.1	0.83	3.9	0.36	57.5	0.98	6.4	0.54	65.1	0.87	2.6	0.32
Digital Reading												
Low	52.0	1.06	8.9	0.60	36.7	1.26	17.7	1.15	61.3	1.04	5.6	0.49
Medium	58.8	1.17	6.1	0.59	43.0	1.26	11.6	0.92	68.2	1.23	4.0	0.56
High	65.1	1.19	4.2	0.43	50.2	1.43	9.0	0.85	74.1	1.23	2.4	0.36

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

Table 8 shows percent correct and missing by domain and item type for students in schools in the SSP under DEIS and students not in these schools. Across all domains, percent missing is lower and percent correct is higher for students in non-SSP schools, and this is true of both written response and multiple choice item types. However, the results for print and digital reading are worth comparing. The difference in percent correct for these two groups of students is smaller for digital reading (11.2%) than for print reading (15.0%). There is also a much smaller difference between SSP and non-SSP students in the percentage of missing items on digital reading (9.1% and 5.8% respectively) compared with print reading (15.4% and 6.6% respectively).

[8] Students have been allocated to these three ESCS groups on the basis of the (weighted) ESCS score cut-points associated with the 33rd and 67th percentiles for Ireland.

Table 8: Student percent correct and percent missing, overall and by item type, for print reading, mathematics, and digital reading, PISA 2009 – comparisons by school SSP status (Ireland)

	All Items				Written Items				Multiple Choice Items			
	Correct		Missing		Correct		Missing		Correct		Missing	
Print Reading	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
Not in SSP	61.9	0.69	6.6	0.45	61.7	0.73	9.8	0.61	61.3	0.67	3.3	0.32
In SSP	46.9	1.15	15.4	1.46	45.8	1.26	22.8	1.65	47.0	1.14	8.1	1.45
Mathematics												
Not in SSP	45.9	0.63	9.9	0.44	37.5	0.70	16.4	0.66	56.5	0.71	2.9	0.28
In SSP	35.0	1.07	18.9	1.27	25.8	1.03	30.3	1.49	46.1	1.22	6.8	1.13
Science												
Not in SSP	56.9	0.66	4.8	0.38	52.5	0.79	8.3	0.57	59.8	0.64	3.0	0.31
In SSP	44.1	1.30	11.6	1.68	38.8	1.49	18.8	2.06	47.5	1.38	7.9	1.57
Digital Reading												
Not in SSP	60.4	0.80	5.8	0.35	45.1	0.89	11.4	0.66	69.6	0.84	3.7	0.30
In SSP	51.6	1.32	9.1	0.91	36.1	1.69	18.1	1.46	61.0	1.28	5.7	0.82

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

Table 9 shows item statistics (percent correct and missing on all items, written items and multiple choice items) for students in outlier and non-outlier schools, for print reading, mathematics, science, and digital reading[9]. In the case of print reading, mathematics and science, there are very marked differences in the percentages of correct and missing responses. There are also differences between outlier and non-outlier schools in response patterns on the digital reading assessment, though these are not as marked as for the print assessment.

In print reading, there is a difference of 25.9 percentage points in overall percent correct, and the percentage of missing responses also differs substantially (7.4% missing in non-outlier schools compared to 32.5% in outlier schools). The same pattern holds across item types, with particularly high rates of missing responses (42.6%) in outlier schools on written items. For mathematics, there are again marked differences, and students in outlier schools responded correctly to just 17.7% of written items. Missing responses are much higher for written mathematics items (42.4%) than multiple choice mathematics items (17.0%) in outlier schools. Also in science, the percent correct is much lower and percent missing much higher for students in outlier schools, particularly on written response items.

Response patterns on digital reading contrast quite strongly with those for the print assessment. There is only a 13.2% difference in overall percent correct on digital reading (compared with 25.9% on the print reading assessment), and rates of missingness are also lower for students in outlier schools on the digital reading items (13.4%) compared to the print reading items (32.5%). Percentage of missing responses is again highest for written response items for students in outlier schools on the digital reading assessment (26.2%), though notably lower than that for written print reading items (42.6%).

[9] In previous analyses of the PISA 2009 achievement data, it was noted (Cosgrove et al., 2010) that eight schools in the PISA 2009 sample of 144 schools had very low average reading performance, i.e. 100 points or more below the national average. No such schools were in the PISA 2000 sample. Students in these schools appeared to respond differently to the assessment of digital reading than print reading, even after adjusting for socioeconomic composition and gender differences (Perkins et al., 2012, Chapter 8).

Table 9: Student percent correct and percent missing, overall and by item type, for print reading, mathematics, and digital reading, PISA 2009 – comparisons by school 'outlier' status (Ireland)

	All Items				Written Items				Multiple Choice Items			
	Correct		Missing		Correct		Missing		Correct		Missing	
Print Reading	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
Not outlier	59.8	0.63	7.4	0.36	59.5	0.67	11.3	0.51	59.3	0.62	3.5	0.26
Outlier	33.9	2.37	32.5	3.65	31.7	2.47	42.6	4.06	34.8	2.54	22.6	3.77
Mathematics												
Not outlier	44.3	0.58	10.9	0.40	35.8	0.65	18.4	0.63	55.1	0.63	3.2	0.24
Outlier	25.6	1.44	31.0	3.76	17.7	1.71	42.4	3.06	35.5	1.88	17.0	4.14
Science												
Not outlier	55.1	0.62	5.2	0.30	50.5	0.75	9.3	0.49	58.0	0.60	3.2	0.23
Outlier	31.4	2.73	30.4	4.16	25.3	2.77	41.6	4.14	34.9	3.01	24.9	4.23
Digital Reading												
Not outlier	59.0	0.76	6.2	0.37	43.7	0.88	12.3	0.66	68.1	0.78	3.9	0.32
Outlier	45.8	5.14	13.4	2.53	27.7	5.90	26.2	4.30	57.0	4.73	8.4	2.18

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

Percent Correct and Percent Missing – Multiple Regression Analyses

This section looks at percent correct and percent missing as outcomes in a series of multiple regression analyses, with student gender, grade level, ESCS, school type and school SSP status as explanatory variables. Where appropriate, non-significant explanatory variables have been removed from the analyses. An advantage to this approach is that it offers a straightforward synthesis of the descriptive results provided in the previous section. Also, comparing the percentage of variance explained in percent correct and percent missing can also give an indication of the relative importance of this particular set of variables when considering the achievement results for PISA 2009. For each domain (and later, in the case of print and digital reading, position), six outcomes are examined:

- Percent correct on all items
- Percent missing on all items
- Percent correct on multiple choice items
- Percent missing on multiple choice items
- Percent correct on written response items
- Percent missing on written response items.

The explanatory variables were entered into the multiple regression models[10] as follows:

- Gender: 0=male, 1=female (i.e. the bars in the figures that follow represent the change in percent correct and percent missing in respect of females compared with males)
- ESCS (economic, social and cultural status): mean=0, sd=1 (bars in the figures represent the expected percentage change in correct and missing responses corresponding to a one-standard deviation increase in ESCS)

[10] Single-level regression rather than multilevel modelling to allow for direct comparisons of parameter estimates across school- and student-level variables.

- Grade: dummy variables for Grades 8, 10 and 11, with Grade 9 as the reference group (bars in the figures show the changes in percent correct and percent missing for students in grades 10 and 11 compared with students in grade 9)[11]
- School type: dummy variables for vocational, mixed secondary and community/comprehensive, with single sex secondary as the reference group (i.e. bars in the figures show the change in percent correct and percent missing for vocational, mixed secondary and community comprehensive schools compared with secondary schools)
- SSP status: 0=not in SSP, 1=in SSP (bars in the figure show the change in percent correct or percent missing for students in SSP schools compared to students not in SSP schools).

The first part of this section examines the results for all students (drawing together the information provided in Tables 4 to 8), while the second examines results taking positioning of items into account (and hence involves a reduced pool of students). The proceeding section examines the extent to which 'outlier' schools are relevant to a consideration of variations in the response patterns of students.[12]

Figures 1 to 4 present the results of the regression analyses that were carried out on all students that participated in PISA 2009 for each of the four domains. The figures display the effect sizes (i.e., the standardized Beta coefficients) of the explanatory variables (student gender, ESCS, grade level and school SSP status) on percent missing and percent correct, overall and by item type. School type was initially included in the regressions, but was then dropped as it was found not to exert a significant effect on any of the outcomes for print reading in the presence of the other variables in the model. The results for Grade 8 are not displayed in the figures, due to the small percentage of students at this Grade level (2.4% of all students), though as noted above, they were included in the regression analyses (see Appendix, Table A3 for results for all variables).

The most striking finding when Figures 1 to 4 are compared is the varying impact of school SSP status on the overall percent correct and percent missing in the different domains. School SSP status exerts the greatest (negative) influence on overall percent correct for print reading, and the smallest for digital reading, with science and mathematics in between. The parameter estimates associated with percent correct are similar for written response and multiple choice questions across domains. This is despite the fact that, for digital reading, mathematics and science, written response items are more difficult on average than multiple choice items, whereas in print reading, the difficulty levels of multiple choice and written response items are about the same (see Table 4, above).

The parameter estimates associated with percent missing that are attributable to school SSP status are similar for print reading, mathematics and science, but again much smaller for digital reading. The effect of school SSP status on percent missing is greater for written response than multiple choice items across all across domains.

The effect of being a Grade 10 student on overall percent correct also varies by domain, with the greatest (positive) impact associated with digital reading, followed by print reading. The positive

[11] Grade 8 = Second Year, Grade 9 = Third Year, Grade 10 = Transition Year, Grade 11 = Fifth Year. The results for Grade 8 students have been removed from the figures in this section due to the small percentage (2.4%) in this group; however, Grade 8 students are included in all regression models and the data underlying these is in the Appendix.

[12] 'Outlier' schools are those with unexpectedly low average performance; specifically, 8 of 144 schools that participated in PISA 2009 had average print reading scores 100 points or more below the national average. In previous cycles, no such schools were in the PISA samples.

effect of being a Grade 10 student on overall percent correct for mathematics and science is about half of that for digital reading. This finding may be related to the increased usage of computers during school time by Grade 10 students in some schools.

Figure 1: Effects of student gender, ESCS, grade level and school SSP status on student percent missing and student percent correct for print reading, overall and by item type, for all students: Results of multiple regressions (Ireland)

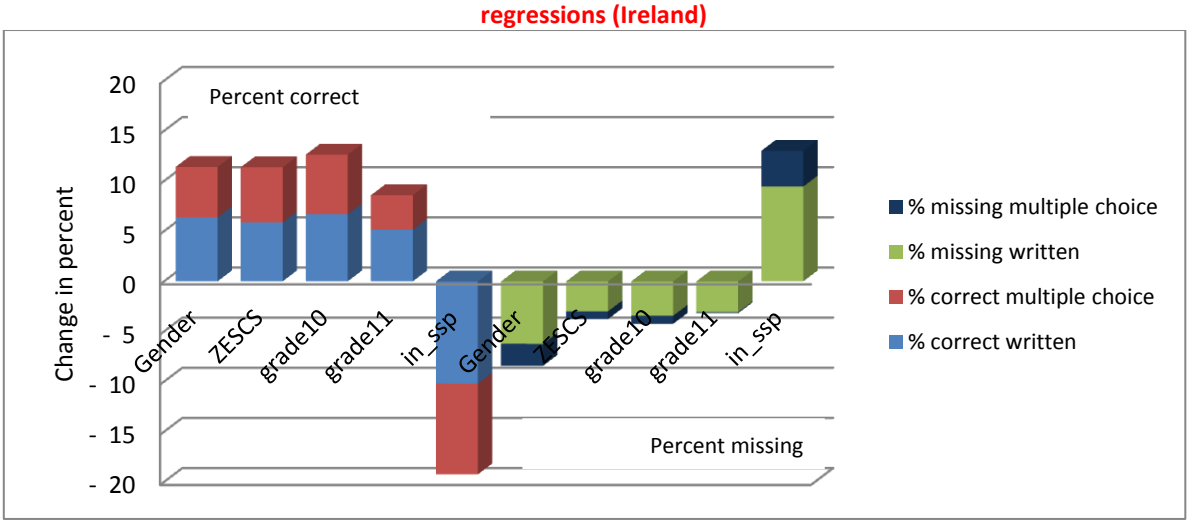


Figure 2: Effects of student gender, ESCS, grade level and school SSP status on student percent missing and student percent correct for digital reading, overall and by item type, for all students: Results of multiple regressions (Ireland)

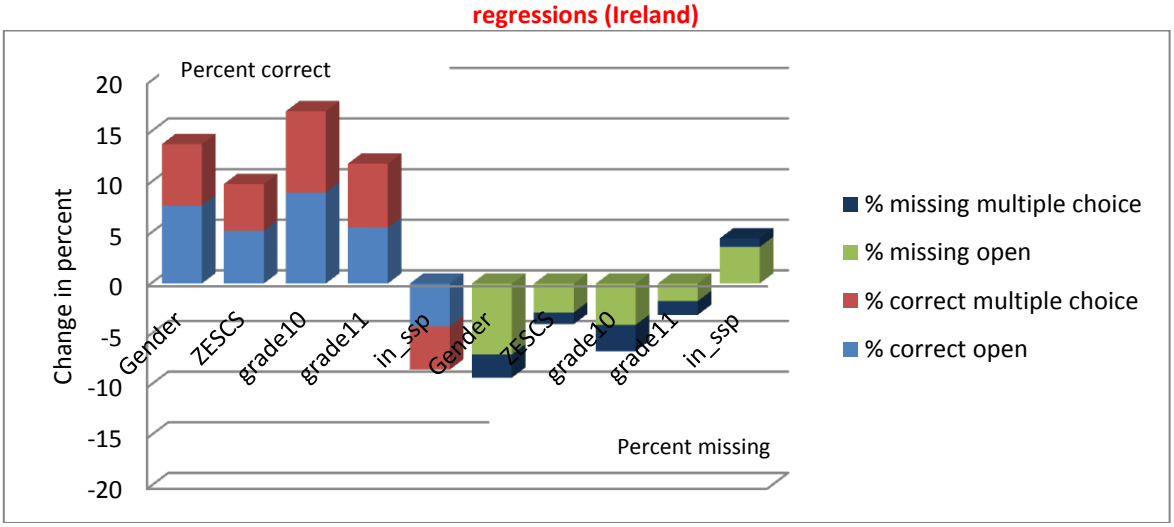


Figure 3: Effects of student gender, ESCS, grade level and school SSP status on student percent missing and student percent correct for mathematics, overall and by item type, for all students: Results of multiple regressions (Ireland)

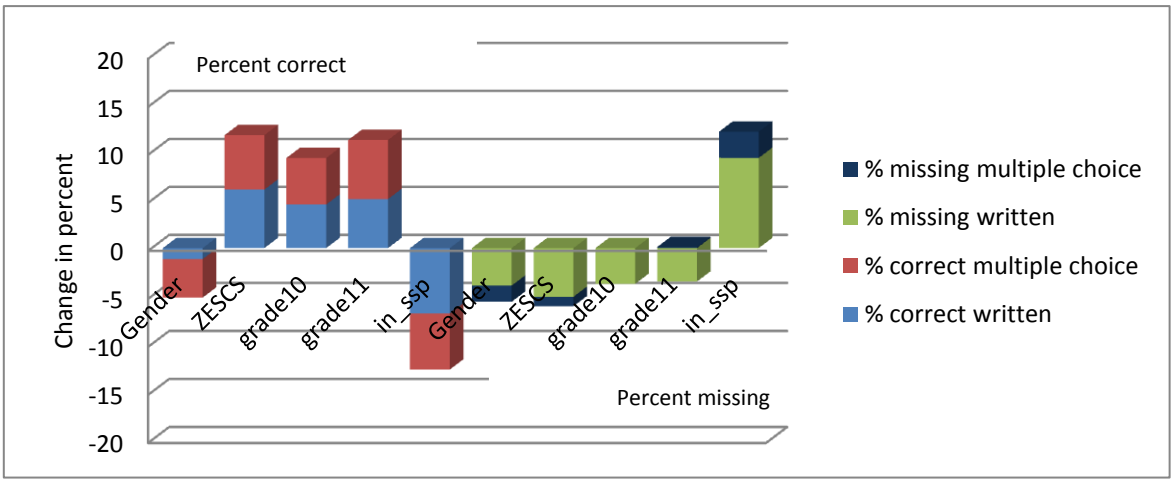


Figure 4: Effects of student gender, ESCS, grade level and school SSP status on student percent missing and student percent correct for science, overall and by item type, for all students: Results of multiple regressions (Ireland)

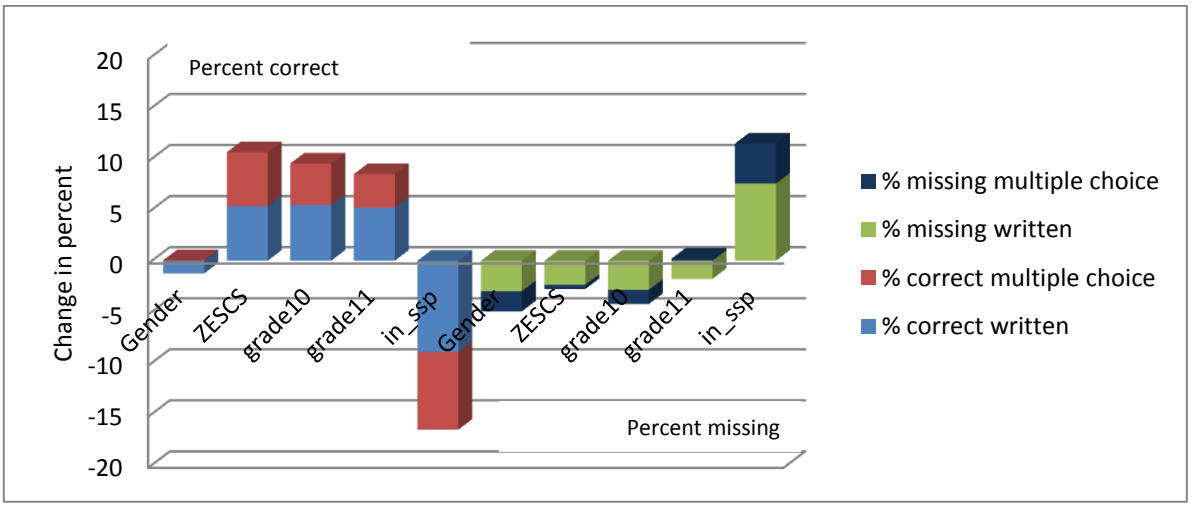


Table 10 displays the variance in each domain/item type combination explained by the regressions displayed in Figures 1 to 4. The explanatory variables account for most variance in overall percent correct in print reading (20.2%), followed by digital reading (16.5%), and finally science (14.7%) and mathematics (14.6%). For missing items overall, the models again explain proportionally more variance in the cases of print reading and mathematics (13.4% in each case) than for mathematics (10.8%) or digital reading (9.4%). It can also be seen that while the models explain similar percentages of variance for correct responses on multiple choice and written response items, the variables in the analyses explain proportionally more of the variance in missing responses on written as compared with multiple choice questions. Comparing Tables 3 and 10, it should be noted that there is no discernible relationship between the explained variance in the percent of missing responses across the domains and the difficulties of items.

Table 10: Percentage of variance in percent correct and percent missing explained by student gender, ESCS, grade level and school SSP status, by item type, for each domain (Ireland)

	% correct	% missing	% correct written	% missing written	% correct multiple choice	% missing multiple choice
Reading	20.2	13.4	17.6	15.2	17.6	5.7
ERA	16.5	9.4	12.6	10.5	14.8	5.5
Mathematics	14.6	13.4	12.5	12.9	10.0	4.6
Science	14.7	10.8	11.3	11.1	13.1	7.5

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

Percent Correct and Percent Missing – Regression Analysis with Position of Items

It was noted earlier that 29.8% of students attempted one of the four booklets containing Block R3A. Since students are randomly assigned to booklets, it can be assumed that the subset of students in the selected schools attempting these particular four booklets is equivalent to all students in this subset of schools. Tables 11 and 12 provide a check of this by comparing all students with those attempting the four booklets in terms of distribution by gender, grade level, school type and SSP status (Table 11), and by average ESCS and performance on print reading, mathematics, science and digital reading (Table 12). Both tables show that the subset of students attempting block R3A are equivalent to the whole sample of students, at least on the characteristics examined.

Table 11: Comparison of all students in Ireland who participated in PISA 2009 and students attempting block R3A, by gender, grade, school type, and SSP status

Group	All Students		Students Attempting Block R3A	
	Weighted N	%	Weighted N	%
Gender				
Male	26732	50.6	7798	49.3
Female	26062	49.4	8029	50.7
Grade level				
Grade 8	1294	2.5	367	2.3
Grade 9	31211	59.1	9248	58.4
Grade 10	12675	24.0	3896	24.6
Grade 11	7614	14.4	2316	14.6
School Type				
Community/Comprehensive	8128	15.4	2299	14.5
Vocational	12199	23.1	3767	23.8
Boys' Secondary	9777	18.5	2840	17.9
Girls' Secondary	11862	22.5	3501	22.1
Mixed Secondary	10828	20.5	3419	21.6
SSP Status				
In SSP	41453	78.5	12384	78.2
Not in SSP	11341	21.5	3443	21.8

Table 12: Comparison of all students in Ireland who participated in PISA 2009 and students attempting block R3A, by average ESCS and mean achievement scores on print reading, mathematics, science, and digital reading

Characteristic	All Students		Students Attempting Block R3A	
	Mean	SD	Mean	SD
ESCS	0.0	1.0	0.0	1.0
Print Reading	495.6	95.1	492.3	95.7
Mathematics	487.1	85.6	487.2	85.4
Science	508.0	97.1	506.2	98.2
Digital Reading	508.9	87.4	508.4	87.4

Table 13 shows the percent correct and percent missing, overall, and by item type, for items in block R3A in each of the four booklet positions, as well as for digital reading percent correct, overall and by item type, for the first and second positions of the assessment. Taking the results for print reading first, it can be seen that, overall, the percentage of correct responses dips markedly in the fourth position, while the percentage of missing responses is also highest in the fourth position. This pattern is apparent across both multiple choice and written response items, but is somewhat more marked in the case of written item types. In contrast, there is little, if any, detectable positioning effect associated with digital reading, except in the case of written response items, where there is a slight increase in the percentage of missing responses in the second position relative to the first, and a marginal decline (from 45.5% to 44.0%) in the percentage of correct responses.

This initial examination of Irish student performance by item position suggests that, in the case of print reading, there is a marked decline in student engagement and/or performance on the test for items presented in the fourth position, particularly on items requiring a written response, while student performance by item position on digital reading is relatively stable by comparison.

Table 13: Percent correct and percent missing for all items, multiple choice items, and written response items, by position: print reading (block R3A) and digital reading (all items) (Ireland)

	All Items				Multiple Choice Items				Written Response Items			
	Correct		Missing		Correct		Missing		Correct		Missing	
Print Reading	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Position 1	53.9	24.7	5.5	10.5	56.9	25.6	2.1	8.5	47.9	29.1	12.2	21.8
Position 2	51.3	24.0	7.3	12.4	54.7	24.3	3.0	10.2	44.4	29.2	15.8	25.6
Position 3	52.2	24.2	9.3	17.7	57.1	24.9	4.4	16.3	42.4	29.6	19.1	28.2
Position 4	45.5	26.2	12.6	23.3	49.5	27.3	7.3	22.2	37.3	29.3	23.2	33.3
Digital Reading												
Position 1	59.5	23.7	4.9	10.2	68.3	24.5	4.1	10.9	45.5	29.9	10.9	22.9
Position 2	59.2	23.1	5.4	10.9	68.4	23.8	4.2	12.7	44.0	30.1	13.8	24.8

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

The remainder of this section explores how these overall patterns in response vary along sub-groups of interest, i.e. student gender, ESCS and grade level, and school sector and SSP status, through a sequence of linear regressions that take percentages correct and missing overall and by item type as the outcome variables. It should be noted that statistical significance of the results cannot be

determined on the basis of the analysis method used (which does not take sampling error into account), so the focus is on general patterns in the results rather than on statistical significance.

Figures 5 to 8 present the results of the regressions carried out on the subset of students who completed the R3A reading block and the digital reading test (see Appendix, Tables A4 and A5 for the data underlying the figures). As noted earlier, this is intended to add to the previous analysis by examining potential positioning effects for print and digital reading. Again, the results for Grade 8 are not displayed in the graphs, due to small numbers of cases, though we have included school sector as the parameter estimates are significant in this case.

The first point to note when comparing the overall percent correct across block position for print reading block R3A (Figure 5) and digital reading (Figure 6) is the stability across block positions on digital by comparison to print reading.

Another difference across domain is that, for digital reading, percent correct tends to increase across positions 1 and 2 (albeit slightly, and with the exception of vocational and mixed secondary students), whereas for R3A it decreases between both positions 1 and 2 and positions 1 and 4. The increase in percent correct for students in SSP schools on the digital reading test is particularly notable: SSP status exerts a significant negative effect on percent correct in position 1, but the effect is no longer significant in position 2. By contrast, the negative effect of SSP status on percent correct increases across the four block positions for R3A.

Considering position effects for R3A alone, we can see that the general tendency is for percent correct for R3A to decrease between positions 1 and 2. The relationships between positioning effects and gender and Grade 11 are atypical, however, with Grade 11 students and female students exhibiting an increase in percent correct between these positions.

It can also be observed, for percent correct on R3A, that patterns diverge when the reading block is in position 3, with some groups exhibiting a worsening of performance (the most notable example being students in vocational schools), and others exhibiting improved performance (e.g., mixed secondary and community/comprehensive students). One possibility is that performance on R3A in position 3 is affected by the preceding unit, which is reading unit R220 (“South Pole”). This unit is, arguably, unappealing for some students of this age group, though there is no way of assessing this empirically.

Figure 5: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent correct for R3A, for all item types, by block position: Results of multiple regressions (Ireland)

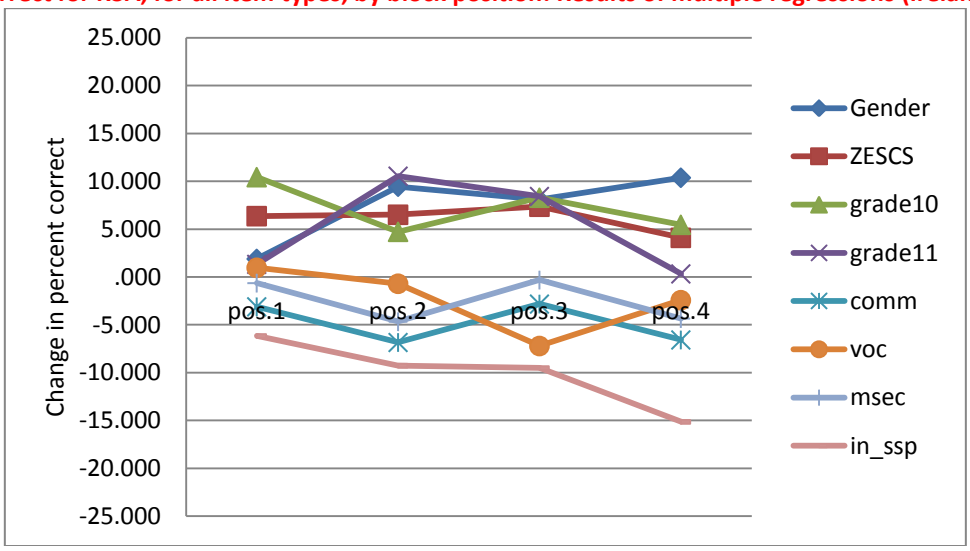
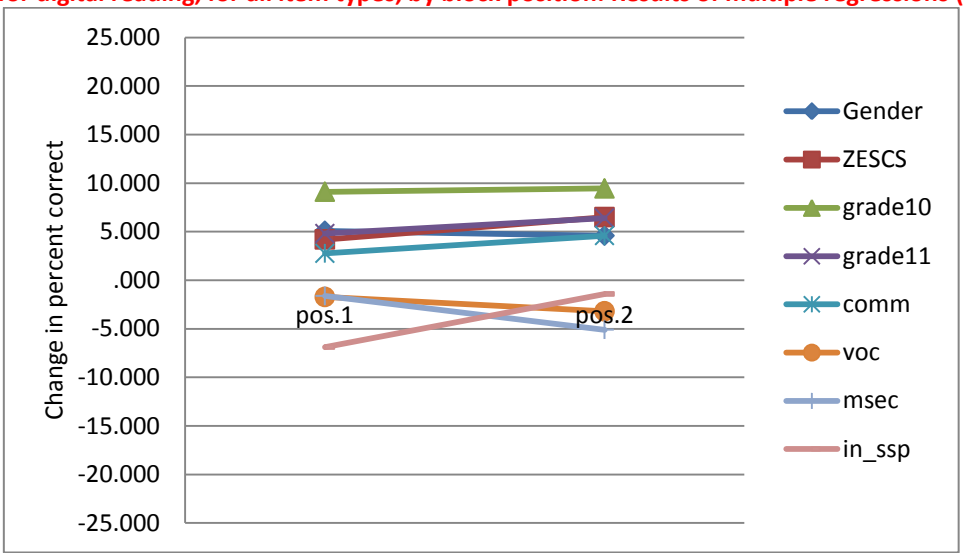


Figure 6: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent correct for digital reading, for all item types, by block position: Results of multiple regressions (Ireland)



Similarly, when percent missing for all items is considered (Figures 7 and 8), there is much less variation across block position for digital reading than for print reading (R3A). For R3A, percent missing tends to show the largest increase in position 4, and the overall increase between positions 1 and 4 is particularly large for students in SSP and community schools. By contrast, the percent missing for SSP students decreases across positions in digital reading.

As with percent correct, percent missing exhibits divergent patterns when R3A is in position 3. The percent missing for R3A is much more stable if positions 1 and 2 are considered than if all block positions are considered, though still not as stable as that seen in digital reading.

Figure 7: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent missing for R3A, for all item types, by block position: Results of multiple regressions (Ireland)

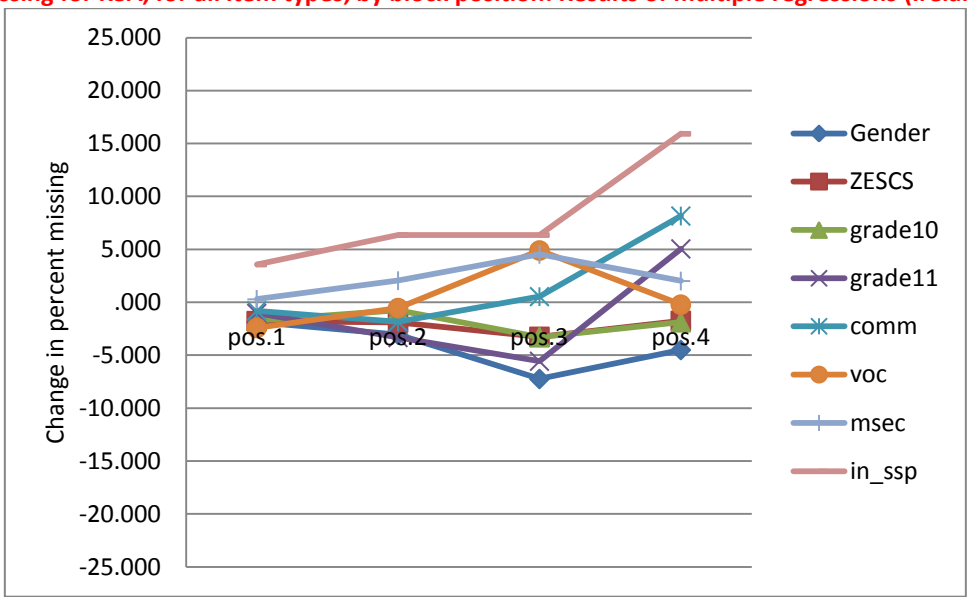
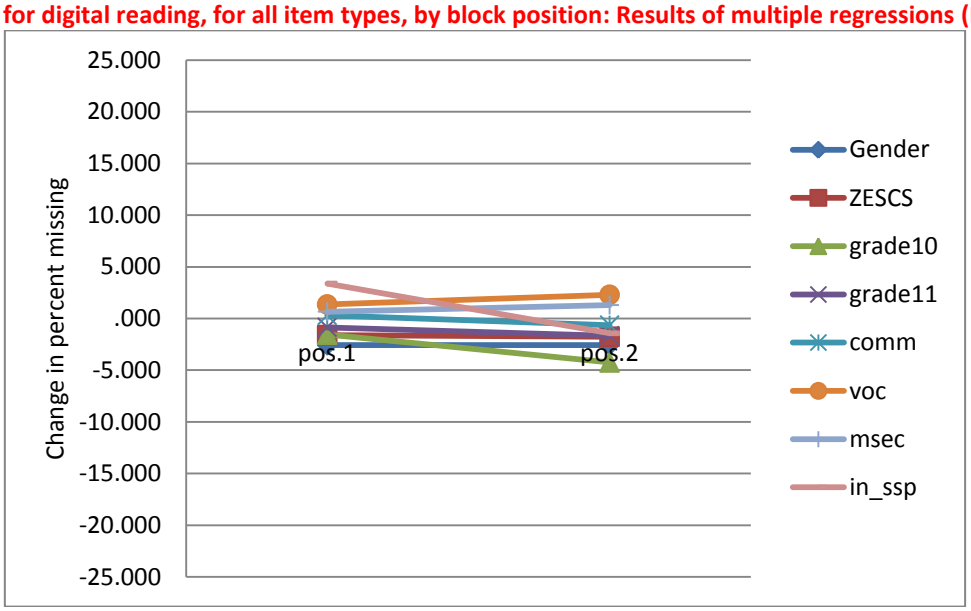


Figure 8: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent missing for digital reading, for all item types, by block position: Results of multiple regressions (Ireland)



As shown in Table 14, the explanatory variables account for more of the variance in outcomes for print reading (R3A) than digital reading, across block positions. The amount of variance in percent correct explained increases slightly across positions 1 and 2 for both R3A and digital reading, but increases sharply in position 4 for R3A (from 29.6% in position 3 to 43.9% in position 4).

Table 14: Percentages of variance in percent correct and percent missing explained by student gender, ESCS, Grade level, school sector/gender composition and school SSP status, for R3A and digital reading (Ireland)

	Position 1		Position 2		Position 3		Position 4	
	% correct	% missing	% correct	% missing	% correct	% missing	% correct	% missing
Print reading (R3A)	23.9	20.8	24.4	25.5	29.6	17.9	43.9	17.8
Digital Reading	17.3	11.2	18.2	11.8	NA	NA	NA	NA

The final set of Figures (9 to 16) presents the same data, but this time it is split by item type (see Appendix, Tables A4 and A5 for the data underlying the figures). Firstly, we focus on differences between item types within print and digital reading, starting with print reading (block R3A; Figures 9 to 12). The most striking difference between item types, for both percent correct and percent missing, is that there is a lot more instability across block position for written response than multiple choice items.

Across positions 1 and 2, there is a tendency for percent correct to increase for written response items, and to decrease for multiple choice items. This is particularly notable for students in vocational schools. An exception to this is gender, with female students showing an increase in percent correct in both written response and multiple choice items. SSP status is also an exception, showing drops in percent correct for both item types across all four positions. Finally, Grade 11 students display a large decrease in percent correct (and an increase in percent missing) in position 4 for the written response items, but not for the multiple choice items.

Figure 9: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent correct for R3A, for written items, by block position: Results of multiple regressions (Ireland)

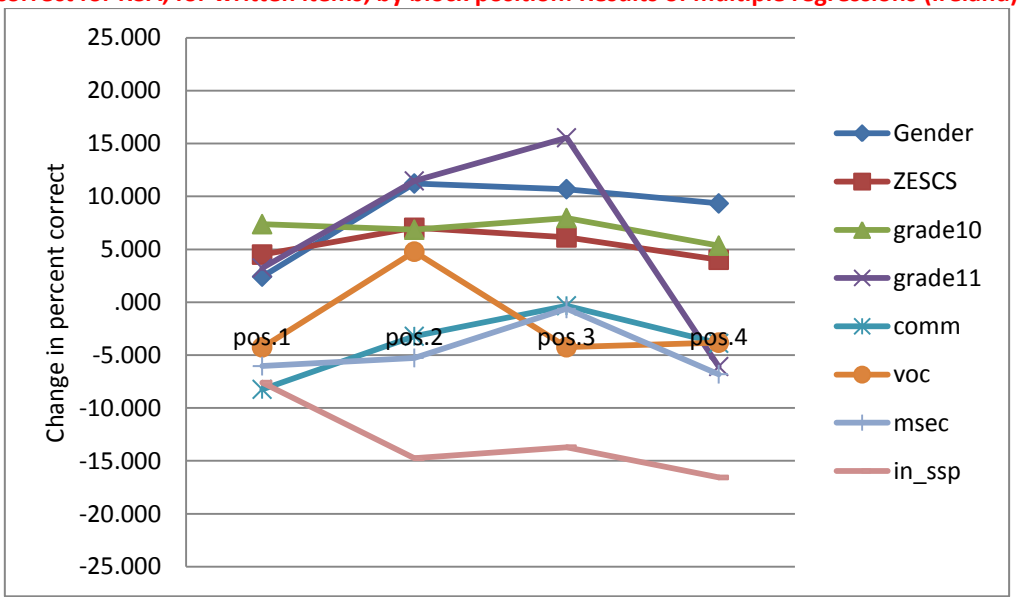


Figure 10: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent correct for R3A, for multiple choice items, by block position: Results of multiple regressions (Ireland)

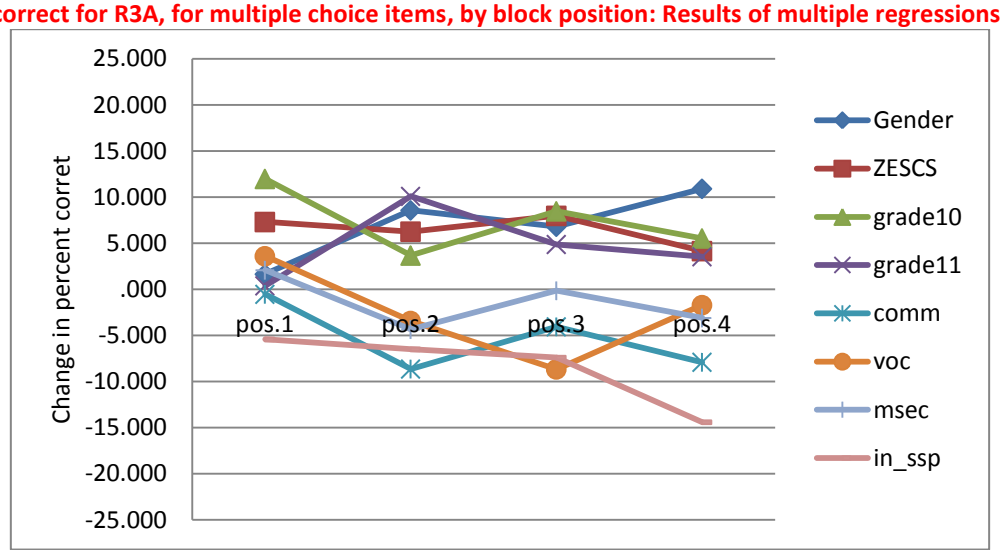


Figure 11: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent missing for R3A, for written items, by block position: Results of multiple regressions (Ireland)

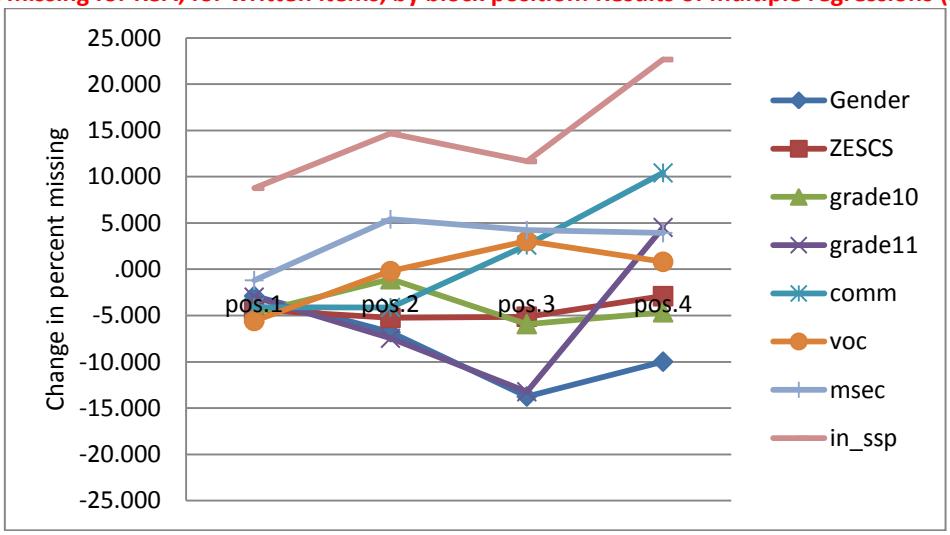
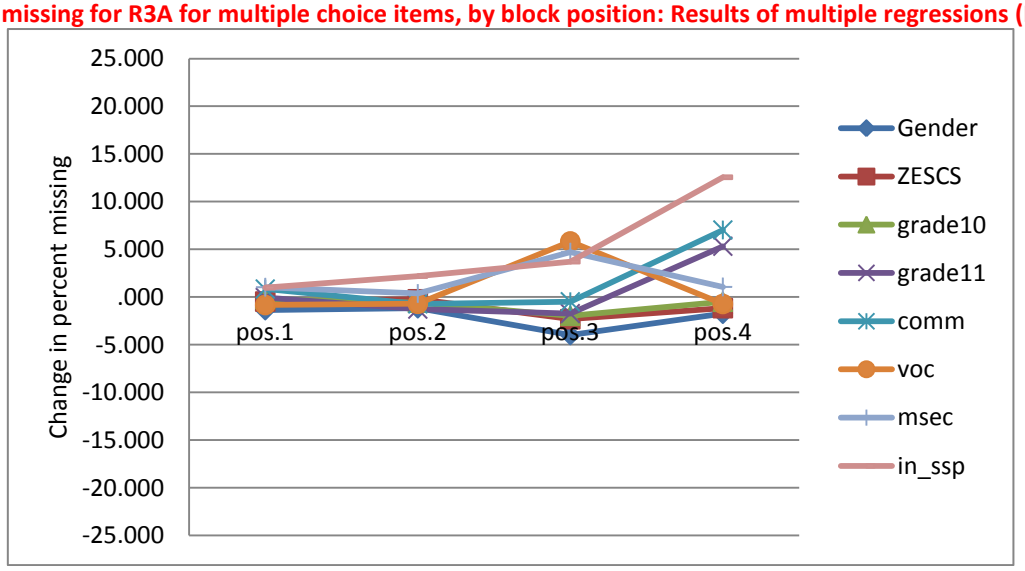


Figure 12: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent missing for R3A for multiple choice items, by block position: Results of multiple regressions (Ireland)



Figures 13 to 16 display the regression results by item type for digital reading. As with print reading, there is less stability across positions for written than multiple choice items, for both percent correct and percent missing. Students in vocational schools show a more marked positioning effect on percent correct by item type, with a decrease in percent correct on written response items from positions 1 to 2 (along with an increase in percent missing) and an increase in percent correct on multiple choice items (with very little change in percent missing). Students in Grade 10 also display an increase in percent correct for multiple choice items and a decrease in percent correct for written response items.

Comparing the results for print and digital reading, the most notable difference is that the effects of SSP status increase across block positions for R3A (for both item types and both outcomes) and decrease across block positions for digital reading (again, for both item types and both outcomes).

Figure 13: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent correct for digital reading, for written items, by position: Results of multiple regressions (Ireland)

Figure 14: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent correct for digital reading, for multiple choice items, by position: Results of multiple regressions (Ireland)

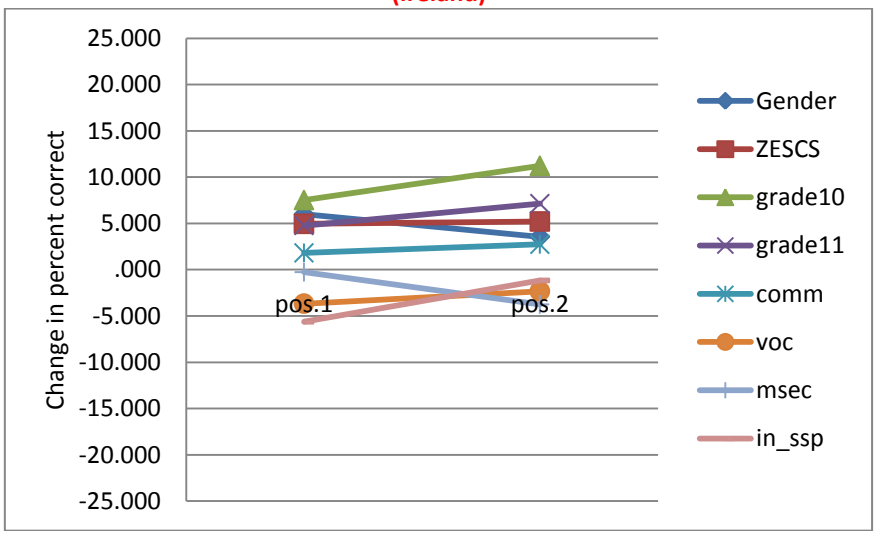


Figure 15: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent missing for digital reading, for written items, by block position: Results of multiple regressions (Ireland)

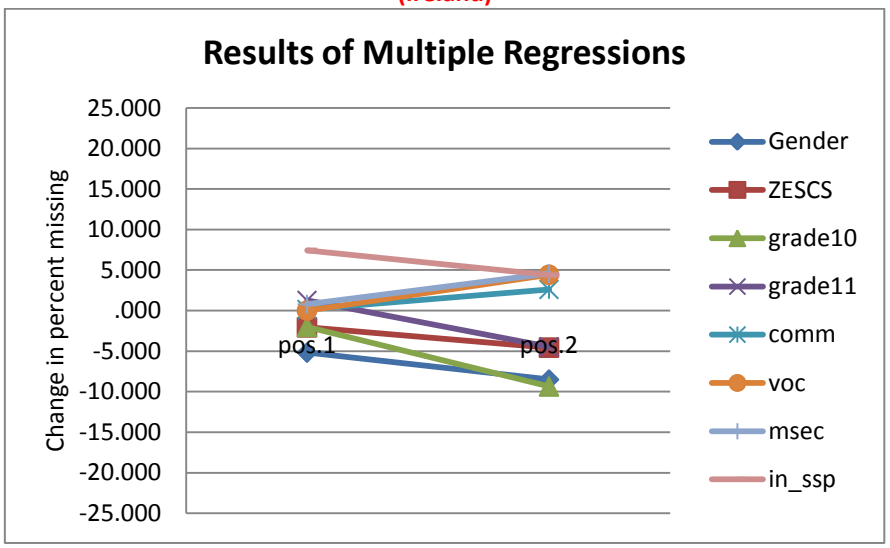


Figure 16: Effects of gender, ESCS, Grade level, school sector/gender and school SSP status on student percent missing for digital reading, for multiple choice items, by block position: Results of multiple regressions (Ireland)

As shown in Table 14, the amount of variance in percent correct accounted for by the explanatory variables for written response items increases between positions 1 and 2, for both print and digital reading. By contrast, the amount of variance in percent correct attributable to the explanatory variables for multiple choice items decreases between positions 1 and 2, for both domains. Given the mixed domain design of the print assessment, and differences in item difficulties by item type as well as the percentages of each type associated with print and digital reading, the explained variances associated with positions 3 and 4 for print reading are difficult to interpret. There may be interactions (for example) with percent correct and percent missing between R3A in positions 3 and 4 that have not been captured by the present analyses. To clarify this, a pooled analysis by position of all seven reading blocks in the print assessment would be required.

Table 14: Percentages of variance in percent correct and percent missing explained by student gender, ESCS, Grade level, school sector/gender composition and school SSP status, by item type, for R3A and digital reading (Ireland)

	Position 1		Position 2		Position 3		Position 4	
	% correct	% missing	% correct	% missing	% correct	% missing	% correct	% missing
Print – Written	16.5	28.5	20.7	20.8	21.4	22.5	21.1	16.4
Print – MC	23.8	2.5	22.0	16.7	28.4	9.9	21.7	14.1
Digital – Written	9.6	9.8	14.0	16.9	NA	NA	NA	NA
Digital –MC	17.0	8.2	14.6	7.1	NA	NA	NA	NA

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

The Relevance of 'Outlier' Schools in Students' Response Patterns in PISA 2009

A further set of multiple regression analyses was carried out, using the same outcome variables, but adding school outlier status to the group of explanatory variables that were considered in the previous section (i.e., student gender, ESCS, grade level, school sector/gender composition and school SSP status). Table 15 presents the effects of school outlier status on percent correct and percent missing for print reading block R3A and digital reading, when the other variables are accounted for.

As expected, school outlier status generally has a negative impact on percent correct and a positive impact on percent missing, across domains, positions and item types (with the exception of percent missing for R3A multiple choice items in position 1, which shows a slight decrease of 1.6% for students in outlier schools, and percent correct for R3A multiple choice items in position 2, which shows a slight increase of 2.4%).

Table 15: Effects of school outlier status on percent correct and percent missing, when student gender, ESCS, Grade level, school sector/gender composition and school SSP status are held constant, by item type, for R3A and digital reading (Ireland)

	Position 1		Position 2		Position 3		Position 4	
	% correct	% missing	% correct	% missing	% correct	% missing	% correct	% missing
Print – Written	-7.0	1.8	-11.6	20.3	-19.6	28.3	-8.8	33.1
Print – MC	-11.8	-1.6	2.4	3.7	-14.9	8.9	-20.3	20.7
Digital – Written	-9.7	6.5	-8.6	11.9	NA	NA	NA	NA
Digital – MC	-5.5	3.5	-4.4	2.2	NA	NA	NA	NA

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

Outlier status has a more pronounced negative impact on some of the outcomes related to the print compared to the digital items. If we first consider percent correct, we can see there is very little change across positions in digital reading, and the negative effects of outlier status actually decrease slightly in position 2, for both item types. Positioning effects on percent correct for print reading are not as straightforward, with a decrease in the effects of outlier status on percent correct for written items in position 4, and for multiple choice items in position 2.

Positioning effects on percent missing show different patterns by item type. For written response items, there is a steady increase in the effects of outlier status on percent missing across positions, for both domains. For multiple choice items, percent missing only increases substantially in position

4 for R3A, while percent missing actually decreases marginally between positions 1 and 2 in digital reading.

Table 16 presents the differences obtained from subtracting the percentages of variance explained by the original set of explanatory variables (without school outlier status) from the percentages of variance explained by the addition of outlier status to the models (see Appendix, Table A6 for the percentages of variance in outcomes explained by the full set of variables including outlier status). Differences are generally small, indicating that outlier status does not have much additional explanatory power. The variance explained by outlier status does increase somewhat, however, in percent missing for both item types in position 4 of the print test (3.6% for written items and 3.1% for multiple choice items), and percent missing for written items in position 3 of the print test (4.6%).

For the majority of the outcomes related to digital reading, the inclusion of outlier status in the model actually decreases the amount of variance explained.

Table 16: Differences in percentages of variance in percent correct and percent missing explained when school outlier status is included in the model, by item type, for R3A and digital reading (Ireland)

	Position 1		Position 2		Position 3		Position 4	
	% correct	% missing	% correct	% missing	% correct	% missing	% correct	% missing
Print – Written	+0.1	0.0	+0.5	+2.1	+2.0	+4.6	+0.3	+3.6
Print – MC	+0.6	+0.1	0.0	+0.4	+1.6	+1.2	+2.0	+3.1
Digital – Written	+0.6	-2.9	-4.5	-6.5	NA	NA	NA	NA
Digital – MC	-4.5	-2.7	-2.9	-3.6	NA	NA	NA	NA

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

Percent Correct and Percent Missing by Position – Relationships with Achievement Scores

Table 15 shows the Pearson's correlations between the percentage of correct and missing responses, overall and by item type and booklet position, and students' PISA scale scores on print and digital reading. Comparing positions 1 and 2 for print and digital reading, it can be seen first that the relationship between percent correct and the scale score is stronger for print than digital reading, and second, that the relationship between percent missing and scale score is stronger in position 1 than 2 in the case of print reading, while the opposite is true in the case of digital reading.

Across positions, and whether one examines all item types or written response and multiple choice items separately, the association between percent missing and scale score is increasing in print reading, while there is no clear pattern in the case of digital reading.

These results indicate that students' response patterns are differentially related to achievement scores in print and digital reading in Ireland by position, though we do not know what these relationships are in the case of other countries.

Table 15: Pearson's correlations between percent correct and percent missing by item type and booklet position and students' achievement scores in print and digital reading (Ireland)

	Print Reading		Digital Reading	
All items	% Correct	% Missing	% Correct	% Missing
Position 1	.864	-.534	.795	-.521
Position 2	.881	-.575	.804	-.442
Position 3	.873	-.602	NA	NA
Position 4	.870	-.643	NA	NA
Multiple choice items				
Position 1	.825	-.253	.729	-.421
Position 2	.851	-.419	.694	-.304
Position 3	.837	-.475	NA	NA
Position 4	.843	-.536	NA	NA
Written items				
Position 1	.748	-.577	.646	-.474
Position 2	.751	-.499	.662	-.483
Position 3	.731	-.584	NA	NA
Position 4	.762	-.633	NA	NA

'Missing' refers to the percentages of students who did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.

Summary and Conclusions

In all, the PISA 2009 assessment consisted of 101 print reading items, 35 mathematics items, 53 science items, and 29 digital reading items. An examination of test item characteristics across domains, item types, and clusters or blocks for the PISA 2009 print and digital assessments indicated that mathematics and print reading contained proportionally more written response items (54% and 52% respectively) than science and digital reading (34% and 24% respectively).

A review of the international item difficulties (item percentages correct) indicated that more print and digital reading items were answered correctly (59% and 60% respectively) compared with science and mathematics (54% and 47% respectively). Furthermore, international item percent correct on print reading items did not vary on average by item type, while for the other three domains, written item types were more difficult than multiple choice ones, particularly in the case of digital reading.

It was also observed that item difficulties varied across clusters, though the amount of cross-cluster or cross-block variation depended on the domain, ranging from 45.3% to 48.3% correct in mathematics, 50.5% to 58.4% correct in science, 55.9% to 66.8% in digital reading, and 52.8% to 68.8% correct in print reading. It should be noted that print reading was assessed using seven clusters while the other three domains were assessed using just three clusters each.

An examination of overall student percent correct by domain and item type for Ireland indicated that average percent correct did not vary by item type for print reading, while for the other three domains, written response items were more difficult. Overall, students in Ireland got 58-59% of print and digital reading items correct, 54% of science items correct, and 44% of mathematics items correct. Some variation in the percentages of missing responses was also apparent, with lower rates

associated with science (6%) and digital reading (7%) than with print reading (9%) and mathematics (12%). Missing response for multiple choice items did not vary across domains, and were in the region of 4%, while missing responses for written response items were highest for mathematics (19%), then print and digital reading (about 13%), and lowest for science (11%).

These percentages of correct and missing responses were then examined across gender/school sector, grade level, ESCS, and school SSP status. Considerable variation was observed, depending on the subgroups considered. Generally, however, both boys and girls in vocational schools did less well in all domains than students in other school types (this does not account for differences in the social intake of school types). Boys in community/comprehensive and vocational schools had the lowest percent correct on print and digital reading items, while girls in vocational schools did least well on mathematics and science items. Broadly speaking, the same pattern was observed across item types. However, girls in community/comprehensive and vocational schools did better on written mathematics items than their overall percent correct on mathematics would suggest. Also, girls in vocational schools did less well on written science items than their overall percent correct on science would suggest.

Comparisons of percent correct by domain and item type across grade levels indicated very low percentages of correct responses, and high rates of missing responses, amongst Grade 8 students (who comprised just 2.4% of the sample), and this was most pronounced for written items in all domains. Percent correct tended to be highest (and percent missing lowest) among Grade 10 students, with slightly lower percentages correct (and higher percentages missing) among Grade 9 and 11 students.

Percent correct and percent missing were also found to vary by student ESCS, with percent correct increasing with ESCS, and percent missing decreasing with ESCS. This pattern was observed across both written and multiple choice items across all domains, though the differences were not as marked for digital reading as for print reading.

Similar patterns were observed when comparing students in SSP schools and those not in SSP schools, and the difference in percent correct on digital reading items (9 percentage points) was again smaller than for print reading (15 percentage points). This was accompanied by a lower rate of percent of missing responses among students in SSP schools on the assessment of digital reading than print reading, particularly on written items.

A comparison of the percentages of correct and missing responses on the print assessment of students in outlier and non-outlier schools indicated very marked differences in response patterns, particularly for written response items. However, differences were much less marked between students in outlier and non-outlier schools on the digital reading assessment. For example, there was a difference in percent correct on print reading of 25.9 percentage points, but only 13.2 percentage points' difference on digital reading.

A series of regression analyses examined the extent to which student gender, grade, ESCS, school sector and SSP status explained variation in percent correct and percent missing by item type (all items, written items, and multiple choice items). In all, 24 regressions were computed (i.e., four domains * six outcomes).

Across all domains and item types, the pattern of correct and missing responses varied in a consistent manner by grade, ESCS and SSP status. That is, with all variables in the models, percent correct on both multiple choice and missing responses was higher for females, students with higher ESCS, and students in Grades 10 and 11 (relative to Grade 9). Percent correct was considerably lower for students in SSP schools, with the exception of digital reading, where the difference was smaller. The magnitude and direction of the effect associated with gender depended on the domain, being most marked (and in favour of females) for print and digital reading.

An examination of the variances explained by these 24 models indicated that overall percent correct was most strongly associated with the explanatory variables in the case of print reading (explaining a little over 20% of the achievement variation). Generally, the explanatory variables explained little of the variation in percent missing on multiple choice items compared with written items.

An examination of the overall student percent correct on print and digital reading by cluster position for those students who attempted block R3A (one of the seven print reading clusters used in PISA 2009) indicated a relatively stable performance on block R3A in positions 1, 2 and 3, with a drop in percent correct in position 4 (i.e. the last quarter of the two-hour booklet attempted by students).

In contrast, there was no appreciable change in percent correct on the digital reading items when positions 1 and 2 were compared (i.e. the first and second halves of the 40-minute digital reading assessment). In both print and digital reading, an increase in the percentage of missing responses on written items could be observed in latter positions, and this increase was particularly marked in the case of position 4 for print reading.

Regressions on percent correct and percent missing for print reading and digital items by position with student gender, grade, ESCS, school sector and SSP status as explanatory variables indicated that variation by position and sub-group was more marked for print than for digital reading. Some variations by subgroup and position for print reading are difficult to interpret (e.g. an increase in percent correct for students in vocational schools for items in position 4; a decrease in percent correct for students in Grade 11 for position 4). Others, however, were perhaps less surprising, such as the relatively steady decrease in percent correct, by about 9 percentage points, for students in SSP schools from positions 1 to 4.

While the regression analyses explained similar proportions of the variance in percent correct and percent missing on digital reading items (regardless of position), these analyses explained increasing percentages of variance in the percent of correct responses on print reading items, with an increase in the explained variance from 24% in position 1 to 44% in position 4. However, for missing responses on print reading, the explained variances were smaller for positions 3 and 4, relative to positions 1 and 2.

Regression analyses that took position into account were also conducted separately for written response and multiple choice items. Overall, the results show that variations in percent correct and percent missing on written response items are much larger in the sub-groups considered for print reading than for digital reading. Also in the case of print reading, the general trend was for an increase in the percentage of non-attempted responses for most sub-groups in position 4 even though the two-hour time limit on the print test was regarded as adequate for most students. Aside from the much lower variation in digital reading across subgroups, positions and item types, perhaps

the most notable difference is that students in SSP schools had a higher percentage of correct responses for items in position 2, which is the opposite of what was found in the case of print reading.

The explained variances for the analyses of student responses by item type and position for print and digital reading are difficult to interpret and it may be the case that modelling relatively few item responses in the case of print reading limits the interpretability of the results. Having said this, the models explained 17% to 21% of variance in percent correct on written responses in the case of print reading, and 22% to 28% of variance in percent correct on multiple choice responses. The only pattern worth noting in the case of the digital reading results in this respect, perhaps, is the increase in explained variance for percent missing on written responses in position 2 relative to position 1.

When these regression analyses were re-computed with school outlier status as an additional explanatory variable it was found that, over and above the other variables in the model, outlier status explained little additional variance in missing and percent correct. However, in positions 3 and 4 of the print reading assessment, the percentage of explained variance of both correct and missing responses increased slightly. This indicates that students in outlier schools tended to disengage more than their peers in non-outlier schools in the latter half of the print reading assessment, after adjusting for gender, ESCS, grade, and school sector and SSP status.

Relationships between students' scale scores and percent correct and percent missing were somewhat weaker in the case of digital reading than print reading. Also, correlations between percent missing on print reading and the scale score in print reading tended to increase from positions 1 to 4. Why this is so is not possible to determine from the present analysis, but it indicates that students of varying ability were somewhat more likely to attempt items on the digital reading assessment than on the print reading assessment and that this tendency varied on the position of items in the assessments.

In conclusion, it is difficult to make strong inferences on the basis of these findings, primarily due to the mixed-domain nature of the PISA test design, and the fact that, in the analyses of positioning effects in particular, smaller numbers of items were involved. However, four key conclusions may be drawn: first, given that domains vary in terms of the average percentage of written response and multiple choice items and their relative difficulties, together with the fact that response patterns by item type and domain vary considerably across sub-groups, indicates that improvements should be made to the test design of PISA in terms of how it balances item type and item difficulty across domains. This point is also made by Cartwright (2011), where he has demonstrated that features of the PISA assessment which should not be related to student performance in the context of measuring trends, notably item format, are in fact strongly related to performance. Cartwright (2011, p. 30) comments that ''the fact that PISA design has such a large influence on student performance in Ireland, especially relative to the influence of schools, suggests that changes in Irish PISA performance over time may be a function more of unintended interactions with the testing situation than with student proficiency in the domains intended to be measured by PISA.''

Second, response patterns on the digital reading assessment differed to the print reading assessment. In an overall sense, response patterns on digital reading were more stable across

positions, item types, and sub-groups than they were for print reading. In fact, there was some evidence that engagement among some students, e.g. those in SSP schools, actually increased during the second half of the digital reading assessment (while the opposite was the case for print reading). Also, percent correct and percent missing on the digital reading assessment tended to be less strongly associated with the background characteristics that we have considered here compared with print reading. This indicates that the digital reading assessment is more equitable, and perhaps a purer measure of student proficiency than the print reading assessment, since responses to digital reading items do not interact with the characteristics of test-takers or the position of the items to the same degree as they do for print reading. Of course, this may change with time, as computer-based testing becomes more the 'norm' in national instructional and assessment contexts as well as in international assessment contexts.

Third, students in outlier schools responded very differently to the digital reading assessment compared with the print assessment. Why this is so cannot be inferred from these analyses. What the results do demonstrate, however, is that very different proficiencies would be attributed to students in outlier schools, depending on which assessment is examined. This finding is consistent with results described in Chapter 8 of the PISA 2009 national report (Perkins et al., 2012) from the multilevel models, namely that after accounting for a range of student characteristics, the expected achievement scores on digital reading in outlier and non-outlier schools did not differ significantly, while there was still a significant difference in the expected achievement scores of these two groups of students on print reading. The regression analyses that took position into account and that included outlier school as an additional variable over student gender, grade, ESCS, and school type and SSP status indicate that students in outlier schools appear to have disengaged from the print assessment more so than students in other schools during the latter half of the PISA test.

Fourth, the pattern of responses on print reading in position 4 provides strong evidence that students are disengaging with the test towards the end, and how this disengagement plays out in terms of percent correct and percent missing across item types and characteristics of the students varies. This suggests that the test is either too long or inappropriately timed. It may be the case that, as test length increases, its capacity to accommodate widely diverging student proficiencies, characteristics, interest levels, etc. decreases. It may also be the case that the mixed domain booklet design is partially contributing to the response patterns observed. Mazzeo and von Davier (2008) view this aspect of the PISA test design as one of the most serious challenges in the estimation of achievement scores, both within an assessment cycle, as well as over time. In any case, our analysis of block position for reading should be carried out for all clusters and all domains to gain a better understanding of the findings presented here, particularly in the context of mixed-domain context effects.

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. (2011). *Does student engagement explain performance on PISA? Comparisons of response patterns on the PISA tests across time.* 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 Educational and Skills.* Dublin: Educational Research Centre.

Department of Education and Science (DES). (2005). *Delivering Equality of Opportunity in Schools (DEIS): An action plan for educational inclusion.* Dublin: Author.

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.

Mazzeo, J., & von Davier, M. (2008). *Review of the Programme for International Student Assessment (PISA) test design: Recommendations for fostering stability in assessment results. Paris: OECD Education Working Papers* (EDU/PISA/GB(2008)28).

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.

Appendix: Additional Data Tables

Table A1: Student percent correct and percent missing, overall and by item type, for print reading, mathematics, science and digital reading, PISA 2009 – comparisons by student gender (Ireland)

	All item types				Written Response				Multiple Choice			
	Correct		Missing		Correct		Missing		Correct		Missing	
Domain	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
Print Reading												
Males	55.5	0.88	10.8	0.72	54.8	0.94	16.0	0.91	55.4	0.87	5.6	0.59
Females	62.0	0.66	6.1	0.39	62.0	0.68	9.1	0.52	61.2	0.68	3.1	0.34
Mathematics												
Males	44.6	0.73	13.3	0.57	35.4	0.78	21.6	0.82	56.0	0.90	4.7	0.45
Females	42.5	0.68	10.2	0.53	34.6	0.83	17.2	0.78	52.5	0.75	2.8	0.35
Science												
Males	53.9	0.88	7.7	0.68	49.6	1.03	12.5	0.86	56.6	0.87	5.3	0.63
Females	54.4	0.75	4.8	0.40	49.4	0.94	8.7	0.63	57.6	0.71	2.9	0.32
Digital Reading												
Males	54.9	0.98	8.4	0.56	39.1	1.12	16.5	1.01	64.5	1.04	5.3	0.49
Females	62.2	0.91	4.6	0.31	47.3	1.19	9.1	0.59	71.0	0.85	2.8	0.28

Table A2: Student percent correct and percent missing, overall and by item type, for print reading, mathematics, science and digital reading, PISA 2009 – comparisons by school type (Ireland)

	All item types				Written Response				Multiple Choice			
	Correct		Missing		Correct		Missing		Correct		Missing	
Domain	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
Print Reading												
Community/ Comprehensive	56.7	1.54	9.4	1.17	56.1	1.79	13.5	1.54	56.6	1.32	5.2	0.91
Vocational	52.6	1.23	12.1	1.38	52.1	1.29	18.0	1.60	52.2	1.29	6.2	1.32
Boys' Secondary	57.9	1.88	9.1	1.32	57.5	1.92	13.6	1.73	57.5	1.87	4.7	0.97
Girls' Secondary	65.1	0.89	4.6	0.37	65.3	0.92	7.0	0.55	64.2	0.96	2.2	0.27
Mixed Secondary	60.7	1.28	7.4	0.82	60.1	1.38	11.1	1.18	60.6	1.27	3.7	0.54
Mathematics												
Community/ Comprehensive	43.1	1.40	12.3	0.91	34.6	1.77	20.0	1.15	54.2	1.47	3.7	0.90
Vocational	39.2	1.25	15.7	1.21	30.6	1.10	24.7	1.47	49.7	1.46	5.9	1.04
Boys' Secondary	45.4	1.29	12.0	0.97	36.4	1.39	20.0	1.48	57.2	1.56	3.7	0.61
Girls' Secondary	44.5	1.04	8.8	0.55	36.2	1.14	15.0	0.90	54.6	1.13	2.4	0.37
Mixed Secondary	46.1	1.10	10.3	1.10	37.7	1.44	17.2	1.67	56.5	0.91	3.0	0.54
Science												
Community/ Comprehensive	52.9	1.57	6.9	1.38	48.3	1.89	11.4	1.80	56.0	1.49	4.7	1.21
Vocational	48.3	1.28	9.4	1.53	43.6	1.38	15.6	1.94	51.2	1.40	6.3	1.42
Boys' Secondary	55.7	1.60	6.3	0.84	50.5	1.83	11.1	1.29	58.9	1.60	3.9	0.67
Girls' Secondary	57.5	1.09	3.6	0.41	53.1	1.35	6.3	0.74	60.3	1.01	2.2	0.29
Mixed Secondary	56.8	1.32	5.1	0.64	52.3	1.63	8.4	0.94	59.6	1.22	3.3	0.52
Digital Reading												
Community/ Comprehensive	60.0	1.47	6.3	0.57	44.6	1.76	13.2	1.12	68.9	1.46	3.8	0.66
Vocational	52.5	1.29	8.7	0.98	37.2	1.42	16.7	1.44	61.9	1.41	5.6	0.87
Boys' Secondary	57.4	2.13	7.2	0.95	41.6	2.42	14.7	1.83	67.2	2.25	4.4	0.69
Girls' Secondary	64.6	1.12	4.0	0.44	49.8	1.63	7.7	0.81	73.3	0.96	2.6	0.37
Mixed Secondary	58.5	1.67	6.1	0.70	42.9	1.77	11.9	1.33	67.9	1.73	3.9	0.62

Table A3: Effects of student gender, ESCS, grade level and school SSP status on student percent missing and student percent correct for print reading, mathematics, science, and digital reading, overall and by item type, for all students: Results of multiple regressions (Ireland)

	All item types		Written Response		Multiple Choice	
Domain	% Correct	% Missing	% Correct	% Missing	% Correct	% Missing
Print Reading						
Gender	**5.8**	**-4.3**	**6.4**	**-6.2**	**5.0**	**-2.2**
ZESCS	**5.7**	**-1.9**	**5.9**	**-3.0**	**5.5**	**-0.8**
Grade 8	**-18.2**	**12.3**	**-18.0**	**15.1**	**-17.4**	**8.8**
Grade 10	**6.3**	**-2.2**	**6.7**	**-3.4**	**5.9**	-0.8
Grade 11	**4.3**	**-1.5**	**5.2**	**-3.1**	**3.4**	-0.1
SSP School	**-9.5**	**6.4**	**-10.2**	**9.5**	**-9.1**	**3.6**
Mathematics						
Gender	**-2.5**	**-2.8**	-1.1	**-3.9**	**-4.0**	**-1.7**
ZESCS	**5.7**	**-2.9**	**6.1**	**-5.1**	**5.7**	**-1.0**
Grade 8	**-13.8**	**13.5**	**-12.4**	**17.4**	**-15.8**	**8.2**
Grade 10	**4.9**	**-2.1**	**4.5**	**-3.7**	**4.8**	0.0
Grade 11	**5.5**	**-1.7**	**5.1**	**-3.5**	**6.2**	0.1
SSP School	**-6.3**	**6.3**	**-6.8**	**9.4**	**-5.8**	**2.7**
Science						
Gender	-0.4	**-2.3**	-1.2	**-3.0**	**0.1**	**-2.0**
ZESCS	**5.3**	**-1.1**	**5.4**	**-2.3**	**5.3**	**-0.4**
Grade 8	**-14.9**	**10.3**	**-17.1**	**16.2**	**-13.6**	**7.5**
Grade 10	**4.6**	**-1.9**	**5.5**	**-2.8**	**4.1**	**-1.4**
Grade 11	**3.8**	-0.4	**5.2**	**-1.8**	**3.3**	0.2
SSP School	**-8.1**	**5.1**	**-8.9**	**7.6**	**-7.6**	**3.9**
Digital Reading						
Gender	**6.8**	**-3.6**	**7.7**	**-7.0**	**6.1**	**-2.3**
ZESCS	**4.9**	**-1.6**	5.2	**-2.9**	**4.6**	**-1.2**
Grade 8	**-17.2**	**8.1**	**-16.2**	**17.5**	**-17.9**	4.4
Grade 10	**8.4**	**-3.0**	**9.0**	**-4.1**	**8.0**	**-2.6**
Grade 11	**6.0**	**-1.5**	**5.5**	-1.7	**6.3**	-1.4
SSP School	**-4.3**	**1.6**	**-4.2**	**3.6**	**-4.3**	0.8

Changes are in percent. Significant effects are in bold.

Table A4: Effects of student gender, ESCS, grade level, school sector/gender composition and school SSP status on student percent missing and student percent correct for print reading block R3A, overall and by item type and block position: Results of multiple regressions (Ireland)

	Position 1		Position 2		Position 3		Position 4	
	% Correct	% Missing	% Correct	% Missing	% Correct	% Missing	% Correct	% Missing
All Item Types								
Gender	**1.9**	**-1.9**	**9.4**	**-3.0**	**8.1**	**-7.2**	**10.4**	**-4.5**
ZESCS	**6.4**	**-1.8**	**6.5**	**-1.9**	**7.4**	**-3.3**	**4.1**	**-1.7**
Grade 8	**-28.0**	**12.3**	**-19.7**	**21.3**	**-31.2**	**17.2**	**-17.5**	**-8.9**
Grade 10	**10.4**	**-1.7**	**4.7**	-.7	**8.3**	**-3.3**	**5.5**	**-1.9**
Grade 11	1.3	**-1.0**	**10.5**	**-3.3**	**8.4**	**-5.6**	.3	**5.1**
Community/Comprehensive	**-3.1**	-.8	**-6.8**	**-1.9**	**-2.8**	.5	**-6.6**	**8.2**
Vocational	1.0	**-2.4**	-.7	-.5	**-7.2**	**4.9**	**-2.4**	-.2
Mixed Secondary	-.6	.3	**-4.7**	**2.1**	-.3	**4.5**	**-4.3**	**2.0**
SSP School	**-6.2**	**3.6**	**-9.3**	**6.4**	**-9.5**	**6.4**	**-15.1**	**15.9**
Written Response Items								
Gender	**2.4**	**-2.9**	**11.2**	**-6.8**	**10.7**	**-13.7**	**9.3**	**-10.0**
ZESCS	**4.5**	**-4.4**	**7.0**	**-5.2**	**6.1**	**-5.1**	**4.0**	**-2.9**
Grade 8	**-29.7**	**33.0**	**-13.3**	**23.5**	**-27.6**	**46.0**	**-13.5**	-4.8
Grade 10	**7.4**	**-4.7**	**6.9**	-1.1	**8.0**	**-5.9**	**5.4**	**-4.6**
Grade 11	**3.2**	**-3.0**	**11.5**	**-7.5**	**15.5**	**-13.2**	**-6.1**	**4.5**
Community/Comprehensive	**-8.2**	**-4.1**	**-3.2**	**-4.1**	-.3	**2.6**	**-3.9**	**10.4**
Vocational	**-4.3**	**-5.6**	**4.8**	-.2	**-4.3**	**3.0**	**-3.8**	.8
Mixed Secondary	**-6.0**	-1.2	**-5.3**	**5.4**	-.6	**4.2**	**-6.8**	**3.9**
SSP School	**-7.6**	**8.8**	**-14.8**	**14.7**	**-13.7**	**11.7**	**-16.6**	**22.7**
Multiple Choice Items								
Gender	**1.6**	**-1.4**	**8.6**	**-1.2**	**6.8**	**-4.0**	**10.9**	**-1.8**
ZESCS	**7.3**	**-.5**	**6.3**	-.2	**8.0**	**-2.3**	**4.1**	**-1.2**
Grade 8	**-27.1**	**1.9**	**-23.0**	**20.2**	**-33.0**	2.9	**-19.5**	**-10.9**
Grade 10	**12.0**	-.2	**3.7**	-.6	**8.5**	**-2.0**	**5.5**	-.5
Grade 11	.4	.0	**10.1**	**-1.3**	**4.9**	**-1.7**	**3.5**	**5.3**
Community/Comprehensive	-.5	.8	**-8.6**	-.7	**-4.1**	-.5	**-7.9**	**7.0**
Vocational	**3.6**	**-.8**	**-3.4**	-.7	**-8.7**	**5.8**	-1.7	-.7
Mixed Secondary	**2.1**	**1.0**	**-4.4**	.4	-.2	**4.7**	**-3.1**	1.1
SSP School	**-5.4**	**1.0**	**-6.5**	**2.2**	**-7.4**	**3.7**	**-14.4**	**12.6**

Changes are in percent. Significant effects are in bold.

Table A5: Effects of student gender, ESCS, grade level, school sector/gender composition and school SSP status on student percent missing and student percent correct for digital reading, overall and by item type and block position: Results of multiple regressions (Ireland)

	Position 1		Position 2	
	% Correct	% Missing	% Correct	% Missing
All Item Types				
Gender	**5.1**	**-2.6**	**4.6**	**-2.6**
ZESCS	**4.2**	**-1.6**	**6.5**	**-1.8**
Grade 8	**-25.7**	**7.4**	**-16.2**	**10.0**
Grade 10	**9.1**	**-1.6**	**9.4**	**-4.3**
Grade 11	**4.8**	**-.9**	**6.4**	**-1.7**
Community/Comprehensive	**2.8**	.3	**4.6**	-.6
Vocational	**-1.7**	**1.4**	**-3.2**	**2.3**
Mixed Secondary	**-1.6**	.7	**-5.1**	**1.3**
SSP School	**-6.9**	**3.4**	-1.4	**-1.4**
Written Response Items				
Gender	**4.4**	**-5.2**	**5.7**	**-8.5**
ZESCS	**3.0**	**-2.1**	**7.8**	**-4.6**
Grade 8	**-28.4**	**27.1**	**-17.8**	**26.9**
Grade 10	**10.3**	**-2.0**	**7.7**	**-9.4**
Grade 11	**2.8**	1.3	**5.1**	**-4.4**
Community/Comprehensive	**2.9**	.1	**7.1**	**2.6**
Vocational	1.0	.0	**-4.2**	**4.4**
Mixed Secondary	**-4.0**	.8	**-6.5**	**4.6**
SSP School	**-7.0**	**7.4**	-1.2	**4.4**
Multiple Choice Items				
Gender	**6.0**	**-2.1**	**3.6**	**-1.6**
ZESCS	**4.9**	**-1.7**	**5.2**	**-1.4**
Grade 8	**-24.5**	1.0	**-19.3**	**9.0**
Grade 10	**7.5**	**-1.7**	**11.2**	**-4.1**
Grade 11	**4.8**	**-1.8**	**7.1**	**-1.8**
Community/Comprehensive	1.8	.4	**2.7**	**-1.3**
Vocational	**-3.7**	**2.5**	**-2.3**	**2.2**
Mixed Secondary	-.3	.6	**-3.8**	.8
SSP School	**-5.6**	**2.7**	-1.2	**-3.9**

Changes are in percent. Significant effects are in bold.

Table A6: Percentages of variance in percent correct and percent missing explained by student gender, ESCS, Grade level, school sector/gender composition, school SSP status and outlier status, by item type, for R3A and digital reading (Ireland)

	Position 1		Position 2		Position 3		Position 4	
	% correct	% missing	% correct	% missing	% correct	% missing	% correct	% missing
Print – Written	16.6	28.5	21.2	22.9	23.4	27.1	21.4	20.0
Print – MC	24.4	2.6	22.0	17.1	30.0	11.1	23.7	17.2
Digital – Written	10.2	6.9	9.5	10.4	NA	NA	NA	NA
Digital – MC	12.5	5.5	11.7	3.5	NA	NA	NA	NA

'Missing' refers to the percentages of students that did not respond to a question, whether or not that item was followed by a valid response to another item in the booklet they attempted.