The Irish Journal of Education, 2002, *xxxiii,* pp. 71-98.

PISA READING LITERACY IN IRELAND: AN EXPANDED MODEL EXPLORING ATTRIBUTES OF SELF-REGULATED LEARNING

Nick Sofroniou, Gerry Shiel, and Judith Cosgrove
Educational Research Centre
St Patrick's College, Dublin

In the national report on the performance of Irish students in the OECD Programme for International Student Assessment (PISA), a model of performance on reading literacy was presented. In this paper, the initial model is expanded using measures of a number of attributes of self-regulated learning, including variables dealing with learning strategies, motivational preferences and volition, and self-concept/self-efficacy. The expanded model, which explains 78.2% of between-school variance and 47.1% of within-school variance, includes several school- and student-level variables found in the earlier model, as well as four student-level attributes of self-regulated learning: instrumental motivation, academic self-concept, and preferences for competitive and co-operative learning. Unlike the initial model, there are no significant main effects or interactions involving gender. Similarities and differences between the initial and expanded models are discussed, as is the absence of such variables as cognitive strategies and control strategies in the expanded model. Implications for policy are considered.

In an initial analysis of the performance of Irish students in the OECD PISA 2000 study, a multi-level model examined the effects of a range of school- and student-level variables on students' performance in reading literacy (Shiel, Cosgrove, Sofroniou, & Kelly, 2001). School-level variables in the model include designated disadvantaged status, school type, and school disciplinary climate, while student-level variables include socioeconomic status, family structure, frequency of absence from school, frequency of completing homework on time, dropout risk, and an interaction between gender and number of books in the home. The initial model also includes some variables specifically associated with reading, including frequency of leisure reading and attitude to reading. The purpose of the study described in the current paper was to expand the model with a number of variables associated with self-regulated learning available in the released PISA database. Research on self-regulated learning is briefly reviewed and the main concepts are defined. Following that, the OECD Programme for International Student Assessment (PISA) is described and its assessment of self-regulated learning is outlined. Finally, issues arising from the development of the initial model of performance on PISA reading literacy are considered.

RESEARCH ON SELF-REGULATED LEARNING

In recent years, there has been considerable interest among researchers and practitioners in how students can be supported in developing self-regulated learning skills. Self-regulated learning has been defined as 'having the ability to develop knowledge, skills and attitudes which enhance and facilitate learning, and which – abstracted from the original learning context – can be transferred to other learning situations' (German PISA Consortium, 2000, p. 1), and is viewed as a 'goal-orientated process of active and constructive knowledge acquisition, involving the guided interaction of an individual's cognitive and motivational/emotional resources' (p. 1).

Discussion on self-regulated learning in the research literature focuses on the importance of cognitive, motivational/volitional, and metacognitive (evaluative) elements of learning. According to Zimmerman (1989), 'students can be described as self-regulated to the degree that they are metacognitively, motivationally, and behaviorally active participants in the learning process' (p. 4). Boekaerts (1997) describes self-regulated learning as a complex, interactive process involving motivational as well as cognitive self-regulated learning. She suggests that self-regulated learning is multi-layered, involving regulation of the self (choice of goals and resources), regulation of the learning process (use of metacognitive knowledge and skills to direct one's learning) and regulation of information processing modes (choice of information processing strategies), and that there is a mutual inter-dependency among components – i.e., deficits in one area are viewed as inhibiting or impeding other aspects of self-regulated learning (Boekaerts, 1999). Baumert (1999) draws attention to another component of self-regulated learning – domain-specific knowledge.

Self-regulated learning skills are viewed by some authors as operating along a developmental trajectory, becoming gradually more differentiated as the learner becomes more aware of the value of strategy use (Borkowski, Milstead, & Hale, 1988; Pressley, Borkowski, & Schneider, 1989). Parallel to this, the motivational components of self-regulated learning, including attribution style, self-concept, and self-efficacy beliefs may also develop and change as the student matures as a learner. Others, with a more behaviourist orientation, emphasize the role of social and physical environments in selecting successful self-regulated learning strategies with reference to their consequences (Catania, 1997; Donohue & Palmer, 1994).

There is support in the literature for individual components of self-regulated learning, including those assessed in PISA (see below). According to Zimmerman and Martinez-Pons (1989), the learning strategies that facilitate self-regulated learning include rehearsing and memorizing, goal-setting and

planning, self-evaluating and self-monitoring, and purposively seeking information. Similarly, Pressley et al. (1989) view competent thinking as a function of knowledge about strategies for achieving goals, knowledge of where and when strategies can be implemented (metacognitive knowledge) and application of these strategies in combination with a comprehensive network of content knowledge. While cognitive strategies include memorization, elaboration, and transformation (transfer of information to another medium), metacognitive strategies include planning (defining goals and formulating control questions), monitoring (checking that material has been understood), and regulation (adapting learning activity to given tasks).

As already suggested, motivation to learn is viewed by some theorists as an attribute of self-regulated learning. In Deci and Ryan's (1985) self-determination theory of learning, a distinction is made between intrinsic motivation (defined as interest in a subject area) and instrumental (extrinsic) motivation. According to this theory, subject-related interest (e.g., interest in reading) affects the continuity and intensity of engagement in learning situations and depth of understanding. This can be viewed as independent of the influence of generalized motivation to learn (German PISA Consortium, 2000). Dimensions of interest include the individual evaluation of content. For Deci and Ryan, the experience of competence, self-efficacy, autonomy, and social integration are necessary for motivational self-regulation.

The literature on self-regulated learning also provides support for the view that self-efficacy beliefs and self-concepts of personal ability are associated with goal setting, strategy use, and academic achievement (Zimmerman, 1999).

The literature on reading development and the teaching of reading provides strong support for the importance of self-regulated learning. For example, Hacker (2000) uses the term 'self-regulated comprehension' to describe monitoring and control of reading processes, and discusses how readers construct and compare cognitive and metacognitive models of text during comprehension. Though not an integral part of his model of comprehension monitoring, he notes that some readers may lack the motivation to control their reading.

Guthrie and Wigfield (2000) focus on the concept of engaged readers, defined as those who 'co-ordinate their strategies and knowledge (cognition) within a community of literacy (social) in order to fulfil their personal goals, desires and intentions (motivational)' (p. 404). They note that such readers can provide themselves with self-generated learning opportunities that are equivalent to several years of education. In this view, engaged readers can overcome obstacles to achievement, and become 'agents of their own reading growth'. Although claiming that motivation, defined as 'the individual's

personal goals, values and beliefs with regard to the topics, processes and outcomes of reading' (p. 405), is distinct from reading attitude and interest in reading, Guthrie and Wigfield note that it is multifaceted, with some individual aspects of motivation more relevant in some situations than in others.

Of particular interest in this regard are training studies which show that instruction in particular aspects of self-regulated learning can have positive effects on students' reading achievement. Programmes that have been found to be successful in enhancing learning in general (e.g., Pintrich & De Groot, 1990) and reading in particular (e.g., Guthrie, Van Meter, McCann, & Wigfield, 1996) serve to underline the importance of identifying aspects of self-regulated learning that may be amenable to instruction.

THE OECD PROGRAMME OF INTERNATIONAL STUDENT ASSESSMENT (PISA)

In 2000, the first cycle of an international assessment involving 15-year olds, PISA, was conducted in 28 OECD member countries (including Ireland) and in four additional countries.[1] In Ireland, a sample of 3,854 15-year olds in 139 schools took part in the assessment. In selecting the sample, schools were chosen first with a probability proportional to their size in each of three strata – large, medium, and small.[2] A fixed number of 15-year olds were then selected at random within each school, yielding a two-stage stratified cluster design.[3] Weighted school and student response rates exceeding 85% were obtained.

Students completed a comprehensive test of reading literacy (the major assessment domain) and less comprehensive tests of mathematical and scientific literacy (the minor domains). In addition to these tests, students completed a questionnaire that sought information about their socioeconomic background, their parents' educational attainment, and their own attitudes towards, and engagement in, reading and learning. The students' principal teachers completed a separate school questionnaire that sought information about school management, organization of learning in schools, and resource availability and their use.

1 One OECD member country, The Netherlands, was not included in several analyses in the PISA 2000 international report because its school-level response rate fell below predefined international standards. The four non-OECD countries are Liechtenstein, the Russian Federation, Latvia, and Brazil.

2 Small schools were defined as those with 17-40 15-year olds, medium as those with 41-80, and large as those with more than 80.

3 35 students (or the actual number of students, if lower) were selected in each school.

The mean score of Irish students on the PISA measure of combined reading literacy was 526.7 (SE = 3.24), while the standard deviation was 93.57 (SE = 1.69). Irish students ranked 5th of 27 OECD countries on this measure, with only one country, Finland, achieving a significantly higher mean score. In addition to performance on combined reading literacy, PISA reported performance on three reading literacy subscales (OECD, 2001; Shiel et al., 2001). The outcome (response) variable in the study described in this paper is the measure of combined reading literacy.

The Assessment of Self-Regulated Learning in PISA

Of particular interest in the context of the current study is a series of 52 items on the PISA student questionnaire that sought to tap into attributes of self-regulated learning.[4] Together, these items represented an international option that was administered in 25 countries including Ireland. Twenty-eight of the items sought responses on a 4-category frequency scale ranging from 'almost never' to 'almost always'. The remaining items asked for students' level of agreement with a series of statements on a 4-category scale ranging from 'disagree' to 'agree'. An exploratory factor analysis was performed by researchers on behalf of the OECD on the international data set following a field trial involving the 25 countries in 1999 to develop a reasonable latent structure for the data. Ten factors with eigenvalues greater than one were identified through principal components analysis and varimax rotation. A pattern of loadings supporting the identification of three aspects of self-regulated learning (learning strategies, motivational preferences, and self-concept) was observed. Following the PISA 2000 study, a factor analysis confirmed the three-aspects solution, while an exploratory factor analysis within each aspect suggested a latent structure of that aspect in terms of the number of subscales and items within subscales. The three aspects and associated subscales are as follows (numbers of items in brackets; see Appendix 1 for more detailed descriptions):

> *learning strategies*, consisting of students' application of memorization/rehearsal (4), elaboration/deep processing (4) and control strategies (5);
> *motivational preferences*, consisting of effort and persistence in learning (4), instrumental motivation (3), competitive learning (4), co-operative

4 According to the German PISA Consortium, the *Motivated Strategies for Learning Questionnaire* (MSLQ) by Pintrich, Smith, Garcia, and McKeachie (1991) and the *Kiel Learning Strategies Inventory* (KLS) by Baumert, Heyn, and Köller (1992) were drawn on in developing the learning strategies component of the self-regulated learning items for PISA 2000.

learning (5), instrumental motivation (mathematics), and interest in reading (3);

self-concept, consisting of control expectations (4), self-efficacy for learning (4), self-concept (verbal) (3), self-concept (academic) (3), and self-concept (mathematics) (3).

In Ireland, reliability estimates (Cronbach's alpha) for subscales ranged from .85 (instrumental motivation) to .75 (memorizing). In the current study, the subscales, with the exception of those tapping self-concept (mathematics) and instrumental motivation (mathematics), were considered as candidate variables for an expanded model of reading literacy.

The Initial Model of PISA Reading Literacy in Ireland

In the PISA study, in which 15-year olds were selected at random within schools, the intra-cluster (school) correlation for combined reading literacy for Ireland was .178, suggesting that between-school variance in reading literacy is relatively small compared with other OECD countries. Indeed, just six countries (Canada, Finland, Iceland, New Zealand, Norway, Sweden) had lower intra-cluster correlations (OECD, 2001, Table 2.4).

The initial model of reading literacy presented in the national report on the performance of Irish students in PISA 2000 explained 77.8% of between-school variance and 44.2% of within-school variance. The model included the following variables:

school-level variables: disciplinary climate, school type (whether secondary, community/comprehensive or vocational), and disadvantaged status (whether designated disadvantaged or not);

student-level variables: gender, socioeconomic status, number of siblings, index of books in the home, dropout risk, frequency of absence from school, frequency of completion of homework on time, current grade level, frequency of leisure reading, attitude to reading (a variable based on frequency of reading, interest in reading, and motivation to read), and an interaction between gender and the index (number) of books in the home.

The model confirmed the associations of a number of variables with achievement, and estimated their contributions to the fitted values for students' scores. For example, attendance at a vocational rather than a community/comprehensive school was estimated to reduce a student's score by 20.4 points, or one-fifth of a standard deviation. Similarly, attendance at a school designated disadvantaged was estimated to reduce a student's score by 22.3 points (about one quarter of a standard deviation), while dropout risk (a binary variable) was estimated to reduce a score by 54.4 points (over one half of a standard deviation).

Variables with positive contributions included socioeconomic status, with the average contribution ranging from +25.9 points (one-quarter of a standard deviation) for students at the mean of the high socioeconomic category (the top third of the distribution of socioeconomic scores) to +3.0 points for students at the mean of the low socioeconomic scores (those in the bottom third). Since the model is additive, it is possible to estimate the contributions of combinations of variables.

In the course of developing the initial model, the parameter associated with the top category of frequency of leisure reading changed substantially when all of the level 1 (student) variables were added to the model. Through a series of comparisons of a model containing frequency of reading only and models containing frequency of leisure reading compared with one other level 1 variable at a time, it emerged that, when attitude to reading was entered with frequency of reading, the parameter for 'more than 60 minutes a day' category changed from +19.4 to –21.2. Subsequent examination indicated that the attitude to reading variable, in addition to asking students about their interest in and motivation for reading, included an indirect measure of frequency of reading. For this reason, we have excluded frequency of reading from analyses.

METHOD

Participants

Participants in the current study consisted of 3,603 15-year olds in Irish post-primary schools who participated in PISA in 2000.[5] As part of PISA, these students took the PISA test of reading literacy and responded to the student questionnaire, including the 52 self-regulated learning items.

Variables

Candidate variables for inclusion in the expanded model of PISA reading literacy were selected with reference to: (i) the variables included in the initial model of PISA reading literacy (Shiel et al., 2001) and (ii) variables associated with attributes of self-regulated learning. Frequency of reading was omitted to avoid collinearity problems with attitude to reading, which had a stronger association with achievement, and was of theoretical interest in the context of

5 A listwise deletion procedure was selected for use in modelling. The procedure results in the omission of any students with missing values on any of the selected student-level variables. Of the 3,854 students who completed the test of reading literacy in PISA 2000, questionnaire data on all the student-level variables relevant to this study were available for 3,603 students. Hence, the listwise deletion resulted in the loss of about 6.51% of cases.

self-regulated learning. All of the other student and school variables in the initial model were candidates for inclusion in the revised model.

Correlations among variables designed to assess attributes of self-regulated learning in PISA, variables associated with attitude to reading, the range of students' reading, and combined reading literacy[6] were examined prior to deciding which explanatory variables to consider for inclusion in the expanded model (see Appendix 1 for a description of variables). To reduce multicollinearity problems, the variable having a weaker correlation with combined reading literacy was dropped when correlations between explanatory variables exceeded .60. For example, the correlation between 'interest in reading' and 'attitude to reading' was .787. Since the correlation between attitude to reading and combined reading literacy (.426) was greater than that between interest and combined reading literacy (.362), attitude to reading was retained as a candidate variable. Among the variables that were dropped at this stage due to collinearity were memory and elaboration (learning strategies), control expectations (self-concept), effort and persistence (motivation), and frequency of borrowing library books. The final set of candidate variables is presented in Table 1.

Table 1
School and Student Variables

Student Variables	School Variables
Gender	School Size Stratum
Socioeconomic Status (Combined Scale)	School Type (Sector)
Number of Siblings	Disadvantaged Status
Index of Number of Books in the Home	Negative Disciplinary Climate
Dropout Risk	
Frequency of Absence from School	
Frequency of Homework Completed	
Current Grade Level	
Control Strategies	
Diversity of reading	
Attitude towards reading	
Self-Efficacy for Learning	
Self-Concept - Verbal	
Self-Concept - Academic	
Instrumental Motivation (General)	
Competitive Learning	
Co-operative Learning	

6 Correlations between the full range of self-regulated learning variables and PISA combined reading literacy scores may be found in Shiel et al. (2001), Table 4.46.

Implementation of Modelling Procedures

An expanded multi-level model of performance on combined reading literacy in PISA 2000 was developed using hierarchical linear modelling. The model incorporates a random component at the level of the cluster (schools) that allows for the variation present across clusters to be taken into account. In addition, one or more random coefficients can be included. A random component at the school level consisting of just a random intercept indicates that the slopes of the fitted parameters are constant, but that they vary in a parallel manner from school to school. Adding a random coefficient for a term in the model indicates that the slope for that explanatory variable also varies across schools. The NMLE software library of Pinheiro and Bates (2000), implemented in the R statistical package, was used for model fitting. This was supplemented by functions written to combine the estimates across plausible values[7] (based on Little & Schenker, 1995).

In the model presented here, full maximum likelihood estimation was used, enabling deviance tests of both fixed and random effects to be carried out. In line with the advice of Aitkin , Francis, and Hinde (in press), sampling weights were not applied, and the design strata were incorporated in the model building process. As in the initial model, uncentred continuous variables were used. Hence, the intercept has the conventional interpretation of ordinary least-squares regression (i.e., the value of the linear predictor when the continuous explanatory variables are set to zero). The development of the model involved the procedures described below.

Candidate variables at the student level (level 1) were first evaluated separately as fixed effects in a random intercept-only model of overall achievement on PISA reading literacy, and then evaluated jointly. Non-significant variables were omitted manually using a backwards elimination strategy (with the exception of gender, to enable later evaluation of any gender interactions, which are of policy interest, and co-operative learning, which was of theoretical interest). Any borderline variables ($p \leq 0.1$) were retained in the model at this stage. Categorical variables with more than two levels were evaluated by omnibus tests of deviance changes, fitting the model with and without the corresponding set of dummy variables.

7 In PISA, multiple imputation is used to obtain reliable indices of student proficiency, known as plausible values. The plausible values generated for each student are picked at random from an estimated ability distribution of students with similar item response patterns and backgrounds, and are intended to provide good estimates of parameters of the student populations (e.g., country mean scores), rather than estimates of individual student achievement.

Following this, each level 2 (school) variable was examined separately by adding it to a random intercept-only model as a fixed effect. The explicit stratifying variable, school size (the number of 15-year olds), which had been used in the sample design, was included as a categorical variable, as were the implicit stratifying variables, school gender composition and school type. The evaluations of these variables served to ensure that the unweighted analysis was not distorted by over-sampling of any particular subgroup. The remaining level 2 variables were then simultaneously added to the random intercept model containing the remaining level 1 variables. Variables for which parameter estimates were not statistically significant were removed sequentially. A stricter criterion ($p<.05$) was applied for retention of those explanatory variables which remained.

Next, two omnibus tests of two-way interactions involving gender within and across levels were conducted by adding each such interaction separately to the model.[8] As neither was significant, the non-significant main effect of gender was removed from the model. Following this, omnibus tests of two-way interactions for the remaining variables (at both school and student levels, as well as across levels) were carried out, again by adding the set of interactions involving each variable separately. Significant interactions for pairs of variables which passed the omnibus tests were then added to the model simultaneously. Again, using a backwards elimination strategy, non-significant interactions were dropped. Curvilinearity was then explored through the addition of orthogonal polynomial terms for continuous explanatory variables. Significant polynomials were retained in the model. Following this, random coefficients for each student-level variable were tested separately by addition to the random intercept version of the model. Statistically significant random coefficients were added and evaluated. Next, orthogonal polynomials were refitted as explicit polynomials to facilitate a direct interpretation of their parameter estimates. Finally, a series of submodel fits was computed to allow the calculation of variance explained by the model.

8 Note that in the Irish national report on PISA 2000, interactions between level 1 variables were examined before the addition of level 2 variables to the model. This sequence differs from that used in the present study. The reason for the change is that more interactions would be examined in the current analysis, and it was felt desirable to remove any borderline main effects before examining them.

RESULTS

The development of the expanded model of performance on the PISA 2000 measure of combined reading literacy is presented, followed by estimates of the effects of example values of variables used in the final model to facilitate interpretation of results.

Expanded Model of PISA Reading Literacy

Prior to testing each level 1 variable separately by addition to the random intercept-only model, a preliminary investigation of the curvilinearity of their relationships with the PISA reading literacy scores suggested that the variable 'index of the number of books in the home' would be well represented by its logarithmic form. Table 2 gives the parameter estimates and tests of statistical significance of the level 1 variables that were initially tested separately. Co-operative learning was not significant, but because of its theoretical interest, was also evaluated as part of the simultaneous fit of the explanatory variables.

When all the variables represented in Table 2 were simultaneously entered into the same random intercept model, the variable 'diversity of reading' (which reflects the frequency with which students read a range of materials) was found not to be statistically significant and was removed. The parameter estimate for competitive learning was positive, and the estimates for co-operative learning and instrumental motivation were now significantly negative. Further explorations showed a change of parameter for co-operative learning from nonsignificant to significant and negative, and for instrumental motivation from positive to negative, when academic self-concept was also present in the model. In subsequent model fits, two further attributes of self-regulated learning – control strategies and self-efficacy for learning – were removed. Although not significant, gender was not removed from the model at this stage as its interactions with other variables had yet to be examined.

Table 3 gives the parameter estimates and tests of statistical significance of the level 2 variables that were initially tested separately. When all school-level variables in Table 3 were added to the model, the parameter estimate for school size (stratum) was found not to be significant and was removed. Application

Table 2
Performance Scores on PISA Reading Literacy All Level 1 Variables Tested as Separate Models by Addition to the Random Intercept-Only Model

	Parameter	*SE*	Test Statistic	*df*	*p*
Gender: Male – Female	–28.121	3.727	t = –7.545	24474	<.001
Socioeconomic Status	1.478	0.094	t = 15.784	2310	<.001
Grade Level			χ^2 = 361.285	3	<.001
≤Grade 8–Grade 10	–138.532	8.798			
Grade 9–Grade 10	–45.074	3.996			
≥Grade 11–Grade 10	–9.396	5.021			
Number of Siblings	–6.507	1.040	t = –6.257	2850	<.001
Log Index of Books in the Home	67.088	3.631	t = 18.476	9209	<.001
Dropout Risk: Yes– No	–95.993	4.046	t = –23.727	1485	<.001
Absence			χ^2 = 58.662	2	<.001
No days–1 or 2 days	11.864	3.246			
3 days or more–1 or 2 days	–25.951	5.660			
Homework on Time			χ^2 = 121.386	3	<.001
Never–Mostly	–50.286	7.158			
Sometimes–Mostly	–26.805	3.635			
Always–Mostly	6.622	3.662			
Diversity of Reading	23.061	1.643	t = 14.032	1525	<.001
Attitude to Reading	36.665	1.392	t = 26.339	342	<.001
Control Strategies	16.812	1.327	t = 12.673	1338	<.001
Instrumental Motivation	3.962	1.280	t = 3.095	406	0.002
Competitive Learning	12.865	1.433	t = 8.977	58	<.001
Co-operative Learning	1.716	1.366	t = 1.257	543	0.209
Self-Efficacy for Learning	18.561	1.272	t = 14.587	1022	<.001
Self-Concept – Academic	25.581	1.345	t = 19.019	516	<.001
Self-Concept – Verbal	10.763	1.355	t = 7.945	401	<.001

χ^2 tests are based on deviance differences.

of the stricter p≤.05 criterion for retention in the model led to the removal of the student-level self-concept (verbal). After establishing that there were no significant interactions between gender and any of the school- or student-level variables, gender was dropped from the model. Following the application of further omnibus tests of two-way interactions, seven interactions were found to be statistically significant when evaluated separately by addition to the model: attitude to reading by absence from school, attitude to reading by the log of the index of number of books in the home, competitive learning by dropout risk, absence from school by dropout risk, absence from school by self-concept

(academic), dropout risk by self-concept (academic), and self-concept (academic) by index of number of books in the home.[9] In subsequent models, dropout risk by academic self-concept, academic self-concept by log of the index of books in the home, attitude to reading by absence from school, competitive learning by dropout risk, and absence from school by academic self-concept were removed when evaluated simultaneously.

Table 3
Performance Scores on PISA Combined Reading Literacy: All Level 2 Variables Tested as Separate Models by Addition to the Intercept-Only Model

	Parameter	SE	Test Statistic	*df*	*p*
Negative Disciplinary Climate	–33.609	9.008	t = –3.731	137	<.001
School Type			$\chi^2 = 51.105$	2	<.001
Secondary–Community/Comp	22.460	8.762			
Vocational–Community/Comp	–32.721	10.145			
Not Designated Disadv–Disadvantaged	47.445	6.910	t = 6.866	137	<.001
School Size (Number of 15–Year Olds)			$\chi^2 = 5.576$	2	0.062
Large–Medium	18.521	8.205			
Small–Medium	3.021	16.404			

χ^2 tests are based on deviance differences.

Following tests for curvilinearity, orthogonal polynomials were added to the model for attitude to reading, instrumental motivation, co-operative learning, and number of siblings.

After random coefficients for each student-level variable were tested separately by addition to the random intercept version of the model, the random coefficient for dropout risk was found to be statistically significant, and was added to the model. Finally, polynomials were refitted as explicit polynomial terms to allow direct interpretation of their parameter estimates. The final model (which is interpreted in the next section) is presented in Table 4.

9 Grade level by competitive learning produced a significant change in deviance (p=.041), but was not added to the model as none of its parameters was more than twice the size of their standard errors.

Table 4
Final Expanded Model of Performance on PISA Combined Reading Literacy

	Parameter	SE	Test Statistic	*df*	*p*
Intercept	461.497	9.222			
Student-Level Variables					
Current Grade Level			$\chi^2 = 294.781$	3	<. 001
Grade 8–Grade 10	–92.585	7.452			
Grade 9–Grade 10	–31.45	3.196			
Grade 11–Grade 10	1.514	3.953			
Attitude to Reading	12.416	5.04			
Attitude to Reading Squared	8.163	3.324			
Instrumental Motivation	–4.309	1.281			
Instrumental Motivation Squared	2.408	0.865	t=2.785	67	.007
Competitive Learning	3.569	1.319	t=2.705	30	.011
Co-operative Learning	–2.952	1.325			
Co-operative Learning Squared	–1.886	0.717	t=–2.631	101	.01
Absence from School					
No days–1 or 2 days	4.068	2.997			
3 days or more–1 or 2 days	–4.696	5.373			
Dropout Risk: Yes– No	–38.852	5.756			
Academic Self-Concept	12.823	1.321	t =9.71	310	<.001
Homework on Time			χ^2 =10.437	3	.013
Never–Mostly	–9.713	6.205			
Sometimes–Mostly	–7.473	2.975			
Always–Mostly	–5.801	3.02			
Socioeconomic Status	0.666	0.079	t =8.448	2269	<.001
Number of Siblings	5.38	2.583			
Number of Siblings Squared	–1.448	0.411	t = –3.521	767	<.001
Log (Index of Books in the Home)	31.515	4.075			
School-Level Variables					
Negative Disciplinary Climate	–13.569	4.933	t = –2.751	134	.007
School Type			χ^2 =24.198	2	<.001
Secondary–Community/Comp	1.389	5.494			
Vocational–Community/Comp	–20.759	6.473			
Not Des. Disadv–Disadvantaged	22.346	4.266	t = 5.239	134	<.001
Interactions					
Attitude × Log Index of Books	8.04	3.278			
Attitude Sqrd × Log Index of Books	–4.978	2.069	t =–2.406	45	.02
Absence from Sch. × Dropout Risk			χ^2 =10.604	2	.005
No Days × Dropout Risk	–22.086	7.62			
3 days or more × Dropout Risk	–18.424	10.448			
Random Components					
	Variance	*SD*	*r(Int)*		
Intercept	211.863	14.556			
Dropout Risk	222.312	14.91	0.684	χ^2=12.216	2	.002
Residual	4087.332	63.932			

Variables Dropped from the Model (in sequence)

Diversity of Reading	Dropout Risk × Academic Self-Concept
Control Strategies	Academic Self Concept × Log Books in the Home
Self Efficacy	Attitude to Reading × Absence from School
School Size (Stratum)	Competitive Learning × Dropout Risk
Self-Concept – Verbal	Absence from School × Academic Self-Concept

χ^2 tests are based on deviance differences.

The intra-cluster correlation of the random intercept-only model (i.e., the model with no explanatory variables), which can be interpreted as the proportion of total variability occurring at the school level, is 0.169. This indicates that 16.9% of the variance in achievement can be attributed to the school level, and the remainder to the student/class level (no information on the assignment of students to classes was gathered in PISA).

To estimate the proportion of variance in student achievement explained by the model at both school and student levels, the variance components associated with the model prior to inclusion of the random coefficients for dropout risk were calculated. Using the mean number of 15-year olds enrolled in the population (86.88) as a representative cluster size and Snijders and Bosker's (1999) formula for calculating two-level R^2 statistics, the level 2 R^2 is 0.782, while the level 1 R^2 is 0.471. Hence, the final model explains 78.2% of the variance in achievement at the school level, and 47.1% at the student/class level. The variance explained by the reduced version of the final model that omits the school-level variables is 59.7% at the cluster (school) level and 43.8% at the student/class level. Therefore, the additional variance explained by including the school-level variables is 18.5% at the cluster (school) level, and 3.4% at the student/class level.

Contribution of Variables to Achievement on Combined Reading Literacy

The contribution of a number of variables to the linear predictor, with example values, is examined in this section. Example values are useful when the individual parameter estimates cannot immediately be translated into units of the response variable (i.e., quadratic fits and variables involved in interactions). To facilitate interpretation, examples of continuous variables fitted as main effects (for example, negative disciplinary climate) are also given, even though their parameters do have a direct interpretation.

The model is additive in the sense that each variable times its parameter estimate makes an added contribution to the linear predictor. The contribution of categorical variables to the linear predictor is immediately apparent from the parameters in the final model (Table 5). For example, it is indicated that a student in a vocational school has a predicted PISA score that is 21 points (one-fifth of a standard deviation) lower than that of a student in a community/comprehensive school (the standard deviation for Irish students on PISA combined reading literacy is 93.6 points); a student in a school that is not designated as disadvantaged (the reference category) is likely to score 22 points (just under one-quarter of a standard deviation) higher than a student in a school designated as disadvantaged; a student in second year (grade 8) is likely to achieve a score

that is 93 points (almost one standard deviation) lower than a student in transition year (grade 10).

Continuous variables were split into high, medium, and low groups, using the values closest to the 33rd and 67th percentiles on their scales as cut-points, and these values were used to estimate the effects of being at the mean of each of these groups, using the parameter values from the model in Table 4. Table 5 gives the contributions for students scoring at the means of the high, medium, and low groups for socioeconomic status. There is a difference of almost 23 points (over one-fifth of a standard deviation) between students in the high and low groups.

Table 5
Contributions to Scores on Combined Reading Literacy Attributable to Socioeconomic Status

Socioeconomic Status	Estimated Contribution to Scores
Low	20.213
Medium	30.949
High	42.627

Fitted values for a continuous variable at the school level, negative disciplinary climate, are given in Table 6. There is an 11-point difference between students at the means of the groups reporting high and low negative disciplinary climate. Though statistically significant, the range of these contributions is small (in standard deviation terms) relative to the values reported for other school-level variables such as school designated disadvantaged status.

Table 6
Contributions to Scores on Combined Reading Literacy Attributable to Negative Disciplinary Climate

Disciplinary Climate	Estimated Contribution to Scores
Low (neg)	6.567
Medium	1.240
High (neg)	–4.025

Table 7 shows the contribution of each of the four attributes of self-regulated learning in the final model to the linear predictor of students' scores on PISA combined reading literacy. The difference between students with low and high levels of instrumental motivation (a variable for which there is also a squared

term) is 12 points (about one-eighth of a standard deviation). While the difference between students with low and medium levels of instrumental motivation is 10 points, that between students with medium and high levels is just 2 points. This indicates that, when controlling for other variables in the model (including attitude to reading, which might be viewed as a measure of intrinsic motivation), students with low levels of instrumental motivation perform somewhat better than students with medium to higher levels.

Table 7
Contributions to Scores on Combined Reading Literacy Attributable to Student Variables

	Estimated Contribution to Scores			
	Instrumental Motivation	Competitive Learning	Co-operative Learning	Academic Self-Concept
Low	10.324	–3.514	1.144	–11.719
Medium	–0.123	0.831	–0.707	5.121
High	–1.629	5.565	–8.545	19.898

Both competitive and co-operative learning were included in the final model. There is an 8-point difference between students reporting high and low preferences for competitive learning. In the case of co-operative learning, there is a 10-point difference between students reporting high and low levels of preference. A high preference for competitive learning makes a positive contribution to achievement in reading literacy, while a high preference for co-operative learning is negatively associated with reading literacy.

In the case of academic self-concept, there is a 32-point difference (just under one-third of a standard deviation) between the contributions to achievement of students reporting high and low levels.

There are two 2-way interactions in the final model – one between attitude to reading and the log of the index of books in the home, the other between frequency of absence from school and dropout risk. Contributions to achievement for different levels of the index of books in the home for students with relatively positive, neutral, and negative attitudes to reading are given in Table 8. Among students with a positive attitude to reading, the difference between those who have zero books in the home and those who have more than 500 is 67 points (over two-thirds of a standard deviation). In contrast, among students with a negative attitude, the difference between those who have zero books at home, and those who have more than 500 is 34 points (just over one-third of a standard deviation).

Table 8
Contributions to Scores on Combined Reading Literacy Attributable to Attitude to Reading X Index of Books in the Home

	Estimated Contribution to Scores		
	Attitude		
Books in the Home	Negative	Neutral	Positive
0	–3.976	–0.866	22.303
1-10	8.050	20.551	46.180
11-50	15.084	33.080	60.147
51-100	20.075	41.969	70.056
101-250	23.947	48.864	77.743
251-500	27.110	54.497	84.023
500+	29.784	59.260	89.333

In the case of the absence from school by dropout risk interaction, there was a detriment of at least 39 points associated with being in the high dropout risk group, in comparison with the low dropout risk group. Within the low risk group, increasing absence from school had a small negative effect. However, in the high risk group, the pattern was more complex, with students in both the low and high frequency of absence groups performing more poorly than those in the middle group (1-2 days absent). It should be noted that the main effect for dropout risk had a random coefficient, which means that this term varies across schools, with a standard deviation of 14.9 points. For a variable with random slope, the range of values that the random slope takes for 95% of the population may be estimated by taking the square root of the variance associated with the random slope (i.e., its standard deviation) and adding ± 1.96 times this value to the parameter estimate. In the case of the dropout risk main effect, the range is –68.1 to –9.6.

Table 9
Contributions to Scores on Combined Reading Literacy Attributable to Absence from School X Dropout Risk

	Estimated Contribution to Scores	
	Dropout Risk	
Absence from School	Low	High
0	4.068	–56.870
1-2 days	0	–38.852
3 or more days	–4.696	–61.972

Finally, the contribution to the literacy performance of students with varying numbers of siblings is given in Table 10. The difference in performance between

students with no siblings and those with 8 or more is 50 points (more than one-half of a standard deviation). As the number of siblings increases beyond two, the contribution to achievement becomes increasingly negative.

Table 10
Contributions to Scores on Combined Reading Literacy Attributable to number of siblings

Number of Siblings	Estimated Contribution to Scores
0	0
2	4.968
4	–1.648
6	–19.848
8	–49.632

CONCLUSION

The purpose of the study described in this paper was to extend the model of PISA reading literacy presented in the first national PISA report (Shiel et al., 2001) by developing a model that included attributes of self-regulated learning in addition to variables included in the initial model. The study was prompted by the importance attached to self-regulated learning in the literature on learning in general, and learning to read in particular. It was also prompted by the prominence given to attributes of self-regulated learning in the initial international report on PISA (OECD, 2001), in which variables such as use of control strategies during learning, use of memorization and elaboration, and preferences for competitive and co-operative learning were regarded as outcomes of schooling in much the same way as reading literacy is considered an outcome, and were not included as explanatory variables in the models of reading literacy presented in the report.

A comparison of the initial and expanded models of PISA reading literacy indicates that whereas the initial model accounted for 77.8% of variance at the school level, and 44.2% of variance at the student/class level, the extended model accounts for 78.2% of variance at the school level and 47.1% at the student/class level. Hence, the proportion of student/class-level variance explained by the extended model represents an improvement of 2.9% over that explained in the initial model, and 0.4% at the school level.

The models are similar to the extent that both include the same three school-level variables: school type, (negative) disciplinary climate, and school designated disadvantaged status. Both models also include student-level variables such as socioeconomic status, number of siblings in a student's family,

and the frequency with which a student completed homework assignments on time. The initial model included attitude to reading, dropout risk, frequency of absence from school, and an interaction between gender and the index of books in the home. The revised model includes an interaction between attitude to reading and the index of books in the home, and between frequency of absence from school and dropout risk (interactions not tested for in the initial model).

The absence of a main effect for gender, or of any interactions between gender and other variables in the extended model is noteworthy. Female students in Ireland achieved a mean combined reading literacy score that was some 29 points (or almost one-third of a standard deviation) higher than the mean of their male counterparts on the PISA combined reading literacy scale. It can be concluded that the variables in the extended model succeed in explaining the gender difference in achievement. What has not been addressed in the current study is exactly which variables account for the difference.

The expanded model can also be compared to a model of Junior Certificate English (JCE) based on the performance of students in PISA who had taken the JCE examination in 1999 or 2000 (Sofroniou, Shiel, & Cosgrove, 2000). The explanatory variables were those used in the initial model of performance on PISA. The JCE model explained 79.2% of between-school variance and 37.3% of within-school variance, and also confirmed the importance of negative disciplinary climate, school type, and school designated disadvantaged status in explaining differences in achievement among students. Unlike the expanded model reported in this paper, however, the JCE model included a student-level parent engagement variable, and a two-way interaction between attitude to reading and gender.

Of particular interest in the present study was the way in which attributes of self-regulated learning might impact on students' reading literacy achievement. First, several of the attributes of self-regulated learning are strongly correlated with one another and with attitude to reading. For example, the correlation between control strategies and memorization (both aspects of learning strategies) is .671 ($p<.001$), while that between control strategies and effort and persistence (an aspect of motivational preferences) is .744 ($p<.001$). Similarly, the correlation between interest in reading (a motivational preference) and attitude to reading is .787 ($p<.001$). Indeed, several attributes of self-regulated learning were eliminated as candidate variables prior to building the model because of their strong correlations with other explanatory variables, and, in some cases, their lower correlations with reading literacy scores. The attitude to reading variable merits particular attention because of its strong association with several other variables. It appears to measure intrinsic motivation and interest in

reading, as well as aspects of frequency of reading. In the present study, the version of this variable that appears in the PISA database was used. The variable is based on students' responses to nine statements about reading. In future work, it will be necessary to examine the components of attitude to reading in more detail.

Second, just four attributes of self-regulated learning remained in the final extended model of reading literacy: instrumental motivation, competitive learning, co-operative learning, and academic self-concept. The fact that several attributes of self-regulated learning, such as interest in reading, memorization, and elaboration strategies, were not selected as candidate variables in the first instance, and that several others, including control strategies, self-efficacy for learning, and verbal self-concept, were removed from the model at a relatively early stage, suggests that at least some of these variables may not have functioned in the way intended in the design of PISA. For example, if variables such as control strategies or self-efficacy had focused on students' reading behaviours rather than on their general learning behaviours, their associations with other variables in the model and with reading literacy might have been different. However, it may also be the case that earlier experimental research did not control for variables associated with control strategies and self-efficacy, and that some of the conclusions that were reached may need to be reviewed.

The interaction between absence from school and dropout risk is complex in that, while contributions to achievement for low dropout risk and frequency of absence from school are small, and move from positive to negative as frequency of absence increases, the fitted contributions for high dropout risk and frequency of absence are negative and large and do not follow an expected pattern. The values with high dropout risk are –57 points for no absences, –39 for 1-2 days absence, and –62 for three or more days absence. Data in Table A2.1 (Appendix 2), which presents descriptive statistics for the cross-tabulation of dropout risk by frequency of absence from school, support the pattern reflected in the model. Those who are at risk of dropping out of school and are never absent from school have a lower mean score on PISA combined reading literacy (432 points) than those who are absent for 1-2 days (443 points). Why students who are absent for 1-2 days outperform students who attend every day is not clear.

The negative parameter estimate for preference for co-operative learning in the final expanded model is also of interest. It will be recalled that, when entered simultaneously into the model with other student-level variables, the sign of its parameter changed from positive to negative in the presence of academic self-concept. To examine this further, descriptive statistics (mean scores across schools) based on a cross-tabulation of academic self-concept and co-operative

learning were computed (Appendix 2, Table A2.2). These show that, while the trend in means for co-operative learning follows that in the model for medium and high levels of academic self-concept, for students with low academic self-concept, the mean reading literacy score for students with a medium preference for co-operative learning (507 points) is higher than the mean score for students with a high (490) or low (486) preference. This serves to underline the complex nature of the relationship between preference for co-operative learning and reading literacy.

How do our findings arising from the expanded model of reading literacy relate to earlier research that has examined relationships between self-efficacy, self-regulated learning, and aspects of performance? Pintrich and De Groot (1990) concluded that, for seventh graders, self-regulated learning (including metacognitive strategies for planning, monitoring, and modifying their cognition) and self-efficacy for learning were strong predicators of academic performance in English classes. Moreover, it was argued that 'the use of self-regulated learning strategies, such as comprehension monitoring, goal setting, planning and effort management and persistence, is essential for classroom performance' (p. 38). Based on the expanded model, it would seem that when a range of other variables are controlled for, control strategies (including metacognitive strategies), control expectations (which include elements of planning), or self-efficacy for learning do not explain performance in reading literacy. It should be noted that Pintrich and De Groot did not include variables such as school type or student socioeconomic status in their regression models. A similar point can be made about studies in the field of reading. For example, Baker and Wigfield (1999), in their analysis of the dimensions of students' motivation to read, and how those dimensions relate to reading achievement, did not examine the effects of such variables as socioeconomic status and ethnicity, except in the context of single-variable tests comparing differences between mean scores.

A number of policy implications arise from a consideration of the expanded model of reading literacy presented in this paper. Some support the conclusions arising from the initial model of performance on PISA reading literacy. The effects of school-level variables such as designated disadvantaged status and school type (sector) on achievement are underlined. Furthermore, the relative importance of variables such as student socioeconomic status and dropout risk (the effect of which varies across schools in both initial and expanded models) is confirmed. The importance of the index of books in the home (a proxy for home educational environment) is also underlined, and, in the present study, involved an interaction with attitude to reading. The importance of completing homework

on time is again supported, though the parameter estimates are considerably smaller than in the initial model.

The finding that the expanded model explains relatively large gender differences in favour of female students also has policy implications to the extent that programmes or strategies to improve the performance of male students will need to take into account the effects of other variables in the expanded model. However, further research will be needed to specify the particular combinations of variables that best explain gender differences in reading literacy.

Both the initial and expanded models of reading raise issues about the definition of variables in the PISA study. It is clear, for example, that the attitude to reading scale, in addition to assessing aspects of interest in reading and intrinsic motivation to read, also assesses frequency of reading indirectly (e.g., 'I cannot sit still and read for more than a few minutes') and that this, in turn, affects its association with variables such as frequency of reading and interest in reading. Indeed, neither frequency of reading nor interest in reading was considered as a candidate variable for the expanded model because of the nature and/or strength of its association with attitude to reading. This underlines the importance of examining the structure of variables such as the attitude to reading variable in greater detail in the future to achieve greater conceptual clarity.

The current study confirms that certain attributes of self-regulated learning, including instrumental motivation, academic self-concept, and competitive learning, are associated with reading literacy. However, other variables that have been shown to be important in the literature on learning and reading (e.g., control strategies and control expectations) were removed from the expanded model as they did not explain performance in reading in the presence of other variables. Moreover, the effects of the self-regulated learning variables that remained in the final model were smaller than their effects when evaluated separately. Given that the research literature on learning strategies and reading strategies calls for instruction in various dimensions of self-regulated learning, including certain cognitive and metacognitive strategies, it is clear that further, more controlled experimental research needs to be carried out before firmer conclusions on the role of instruction in developing these aspects of self-regulated learning can be drawn.

REFERENCES

Aitkin, M., Francis, B. & Hinde, J. (in press). *Statistical modelling in GLIM* (2nd ed.). Oxford: Oxford University Press.

Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly*, *34*, 452-477.

Baumert, J. (1999). *Self-regulated learning as a cross-curricular competence.* Paper presented to the OECD/INES Network A, Paris.

Baumert, J., Heyn, S., & Köller, O. (1992). *Lernstrategien und schulische Leistungen.* (Learning strategies and school performance). Kiel: Institut für die Pädagogik der Naturwissehschaften an der Universität Kiel.

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers and students. *Learning and Instruction, 7,* 161-186.

Boekaerts, M. (1999). Self-regulated learning: Where we are today? *International Journal of Educational Research, 31,* 445-475.

Borkowski, J.G., Milstead, M., & Hale, C. (1988). Components of children's metamemory: Implications for strategy generalization. In F.E. Weinert & M. Perlmutter (Eds.), *Memory development: Universal changes and individual differences* (pp. 73-100). Hillsdale, NJ: Erlbaum.

Catania, A.C. (1997). *Learning* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Deci, E.L., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum Press.

Donohue, J.W., & Palmer, D.C. (1994). *Learning and complex behavior*. Boston: Allyn & Bacon.

German PISA Consortium. (2000). *Self-regulated learning as a cross-curricular competence.* Unpublished report prepared for the OECD PISA Board of Participating Countries.

Guthrie, J.T., & Wigfield, A. (2000). Engagement and motivation in reading. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research (Vol. III,* pp. 403-422). Mahwah, NJ: Erlbaum.

Guthrie, J.T., Van Meter, P., McCann, A.D., & Wigfield, A. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept-orientated reading instruction. *Reading Research Quarterly, 31*, 306-322.

Hacker, D.J. (1998). Self-regulated comprehension during normal reading. In D.J. Hacker, J. Dunlosky, & A.C. Graesser (Eds.), *Metacogniton in educational theory and practice* (pp. 165-191). Mahwah, NJ: Erlbaum.

Little, R., & Schenker, N. (1995). Missing data. In G. A. Arminger, C.C. Clogg, & M.E. Sobel (Eds.), *Handbook of statistical modeling for the social sciences* (pp. 39-69). New York: Plenum Press.

OECD (Organisation for Economic Co-operation and Development). (2001). *Knowledge and skills for life: First results of PISA 2000*. Paris: Author.

Pinheiro, J.C., & Bates, D.M. (2000). *Mixed-effects models in S and S-Plus.* New York: Springer.

Pintrich, P.R., & De Groot, E. V. (1990). Motivational and self-regulatory components of academic performance. *Journal of Educational Psychology*, *82*, 33-40.

Pintrich, P.R., Smith, D.A.F., Garcia, T., & McKeachie, W.J. (1991). *The Motivated Strategies for Learning Questionnaire (MSLQ).* Ann Arbor, MI: NCRIPTAL, University of Michigan.

Pressley, M., Borkowski, J.G., & Schneider, W. (1989). The challenges of classroom strategy instruction. *Elementary School Journal*, *89*, 301-0342.

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

Snijders, A.B., & Bosker, R.J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Sofroniou, N., Shiel, G., & Cosgrove, J. (2000). Explaining performance on the Junior Certificate Examination: A multilevel approach. *Irish Journal of Education, 31,* 25-49.

Zimmerman, B.J., (1989). Models of self-regulated learning and academic achievement. In B.J. Zimmerman & D.H. Schunk (Eds.), *Self-regulated learning and academic achievement. Theory, research and practice* (pp. 1-25). New York: Springer.

Zimmerman, B.J. (1999). Commentary: Towards a cyclically interactive view of self-regulated learning. *International Journal of Educational Research, 31*, 545-551.

Zimmerman, B.J. & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51-59.

Appendix 1

Variables Considered for Inclusion as Candidates in the Expanded Model of PISA Combined Reading Literacy

Variable (Level)	*Description*
Student (Level 1)	
Gender (Male)	Binary (Male, Female); Reference category: Female
Socioeconomic Status (SES)	Continuous - Range: 16 (Min) to 88 (Max);
Number of Siblings (No. Siblings)	Continuous: 0–11; Mode: 2
Index of Number of Books in the Home	Categorical, based on 7 categories: 0, 1-10, 11-50; 51-100; 101-250; 251-500; 501+
Dropout Risk	Binary (High, Low); Based on student's intention to drop out before L. Cert examination; Reference category: No
Absence from School	Categorical: 0 Absences (No absences in two weeks prior to PISA); 1–2 Absences; 3+ Absences; Reference category: 1–2 Absences
Frequency of Completion of Homework on Time	Categorical: Never/Hardly Ever; Few a Year (A few times a year); Once a Month, Several a Month (Several times a month). Reference Category: Never/Hardly Ever
Current Grade Level	Categorical: ≤Grade 8 (Second year or below), Grade 9 (Third year), Grade 10 (Fourth/Transition Year), ≥Grade 11 (Fifth year or higher). Reference Category: Grade 10
Control Strategies	Continuous: Weighted Likelihood estimate (WLE) based on frequency of students' use of five strategies (e.g., 'When I study, I start by figuring out what exactly I need to know'). OECD Mean = 0.0; SD = 1.00.
Diversity of Reading	Continuous: WLE based on frequency with which students read 6 types of texts; OECD Mean = 0.0; SD = 1.00.
Attitude To Reading	Continuous; WLE based on based on student responses to 9 statements on attitude: 'I read only if I have to', 'Reading is one of my favourite hobbies', 'I like talking about books with other people', 'I find it hard to finish books', 'I feel happy if I receive a book as a present', 'For me, reading is a waste of time', 'I enjoy going to a bookstore', 'I read only to get the information I need', and 'I cannot sit still and read for more than a few minutes'. OECD Mean = 0.0; SD = 1.00.
Self-Efficacy for Learning	Continuous: WLE based on frequency with which students indicated that four activities applied to them (e.g., 'I am certain I can understand the most difficult material presented in texts'.) OECD Mean = 0.0; SD = 1.00.
Self-Concept – Verbal	Continuous: WLE based on frequency with which students indicated that three statements applied to them (e.g., 'I learn things quickly in English classes'). OECD Mean = 0.0; SD = 1.00.
Self-Concept – Academic	Continuous: WLE based on frequency with which students indicated that three statements applied to them (e.g., 'I do well in tests in most school subjects'). OECD Mean = 0.0; SD = 1.00

Appendix 1 (cont)

Instrumental Motivation – Gen.	Continuous: WLE based on frequency with which students indicated that each of three statements applied to them (e.g., 'I study to get a good job'). OECD Mean = 0.0; SD = 1.00.
Competitive Learning	Continuous: WLE based on combined frequency with which students indicated that each of 4 statements applied to them (e.g., 'Trying to do better than others makes me work well'). OECD Mean = 0.0; SD = 1.00.
Co-operative Learning	Continuous: WLE based on frequency with which students indicated that each of 5 statements applied to them (e.g., 'I learn the most when I work with other students'). OECD Mean = 0.0; SD = 1.00.
School (Level 2)	
School Size/Stratum	Categorical: Large (81+ 15-year olds); Medium (41–80); Small: 17–40); Reference category: Medium
School Sector	Categorical: Secondary, Community/ Comprehensive, Vocational; Reference category: Vocational
Disadvantaged Status	Categorical (Yes/No); Reference category: Yes
Negative disciplinary Climate	Continuous; Composite/Weighted Likelihood Estimate (WLE) based on student responses to statements about behaviour in class; OECD Mean = 0.0; SD = 1.00
Additional Variables Considered but Not Included as Candidates	
Memorization	Continuous: WLE based on frequency with which students indicated that each of 4 memorisation activities applied to them (e.g., 'When I study, I memorise as much as possible'). OECD Mean = 0.0; SD = 1.00.
Elaboration Strategies	Continuous: WLE based on frequency with which students indicated that each of 4 elaboration strategies applied to them (e.g., 'When I study, I try to understand the material better by relating it to things I already know'). OECD Mean = 0.0; SD = 1.00.
Control Expectations	Continuous: WLE based on frequency with which 4 expectations applied to them (e.g., 'If I want to learn something well, I can'). OECD Mean = 0.0; SD = 1.00.
Interest in Reading	Continuous: WLE based on frequency with which students indicated that each of three statements related to interest in reading applied to them (e.g., 'When I read, I sometimes get totally absorbed'). OECD Mean = 0.0; SD = 1.00.
Frequency of Borrowing Library Books	Categorical: none = no reading; 30 = fewer than 30 minutes per day; 30–60 = 30–60 minutes per day; 60+ = more than 60 minutes
Frequency of Leisure Reading	Categorical variable. Categories: No time; 30 minutes or less per day; 30 minutes to 60 minutes per day; 60 minutes or more per day. OECD Mean = 0.0; SD = 1.00.

Appendix 2

Table A2.1
Interaction Between Absence from School and Dropout Risk – Mean Scores

	Dropout Risk	
Absence	Not at risk	At Risk
No days	549.32	432.48
1- 2 days	536.40	442.78
3 or more days	528.42	406.39

*Weighted estimates based on 5 plausible values

Table A2.2
Interaction Between Academic Self-Concept and Preference For Co-Operative Learning – Mean Scores

	Academic Self-Concept		
Preference for Co-operative Learning	Low	Medium	High
Low	486.45	537.21	574.86
Medium	507.26	528.25	553.90
High	489.73	524.68	542.64

* Weighted estimates based on 5 plausible values