https://doi.org/10.1186/s40536-022-00131-x

Large-scale Assessments in Education

RESEARCH **Open Access**

Examining high achievement in mathematics and science among post-primary students in Ireland: a multilevel binary logistic regression analysis of PISA data

Vasiliki Pitsia[1,2,3*]

*Correspondence:
Vasiliki Pitsia
pitsiavasiliki@gmail.com
[1]Centre for Assessment Research, Policy and Practice in Education (CARPE), Dublin City University, Dublin, Ireland
[2]Educational Research Centre (ERC), Dublin, Ireland
[3]Department of Education, School of Education, University of Nicosia, Nicosia, Cyprus

Abstract
In Ireland, while, on average, students have performed well on national and international assessments of mathematics and science, the low proportions of high achievers in these subjects are noteworthy. Given these patterns and the multifaceted benefits in individual and societal terms that expertise in mathematics and science has been associated with, policymakers in Ireland have begun to place an increasing emphasis on high achievement in these subjects. This emphasis has coincided with ongoing efforts during the last decade to raise interest and improve academic performance within the realm of science, technology, engineering, and mathematics (STEM) education.

Despite this policy attention, research on high achievement in mathematics and science nationally, but also internationally, has been particularly scarce. In an attempt to provide research evidence that could add further impetus to the ongoing efforts, this study examines high achievement in mathematics and science among post-primary students in Ireland using data from the 2012 and 2015 cycles of the Programme for International Student Assessment (PISA). Specifically, the study aimed to evaluate the contribution of various contextual characteristics stemming from students, their families, teachers, and schools in the prediction of high achievement in mathematics and science within a two-stage analysis that included a series of bivariate tests and multilevel binary logistic regression modelling.

The results showed that variables related to students' self-beliefs, engagement, and socioeconomic background were consistently associated with high achievement in mathematics and science. Overall, the significant role of students' homes and families in predicting students' chances of being high achievers in the two subjects was highlighted. In turn, this indicated that further efforts to enhance collaboration between teachers, schools, and parents may be warranted if progress in the area of high achievement in mathematics and science is to be made. The implications of these findings for policy and practice within the Irish context, the limitations of the study, and recommendations for future research are discussed.

Keywords High achievement, Mathematics, Science, PISA, Multilevel binary logistic regression modelling, Ireland

Background

In Ireland, as in many other countries, results from national and international large-scale assessments are often used as key performance indicators for primary and post-primary students, exerting a significant and ever-increasing influence on educational policymaking (Greaney & Kellaghan, 2008; Johansson, 2016). Ireland's results in several national and international large-scale assessments over the last 20 years indicate that while, on average, students perform well in mathematics and science, there are low proportions of high-achieving students in these two subjects and students' scores at the highest national percentiles in these subjects tend to be lower than those of their counterparts in other countries with similar average performance (e.g., McKeown et al., 2019; Perkins & Clerkin, 2020; Pitsia, 2021; Shiel et al., 2016). Such patterns in national and international large-scale assessments have been accompanied by patterns of decreases in the percentages of high achievers in mathematics and science-related subjects in the Junior and Leaving Certificate examinations, the major high-stakes examinations in Ireland (with the latter constituting the university entry examination), over the last 15 years (e.g., Pitsia, 2021; Shiel & Kelleher, 2017). Against this background, students in Ireland have consistently performed very well in reading in both national and international large-scale assessments, with adequate percentages of high achievers in the subject reflective of the respective patterns in average performance.

Within the context of policymakers' heightened interest in science, technology, engineering, and mathematics (STEM) education, high achievement in mathematics and science and the issues described above have attracted considerable attention by the Irish authorities. This is primarily due to the multifaceted benefits to individuals and society with which expertise in these areas, as fundamental components of STEM within compulsory education, has been associated. The *2011 National Strategy to Improve Literacy and Numeracy* was the first governmental policy document to explicitly focus on high-achieving students in mathematics (but not in science, due to the focus of the Strategy on literacy and numeracy only). The Strategy set specific targets for increasing the proportions of students performing at the highest proficiency levels in the National Assessments of Mathematics and English Reading (NAMER) (i.e., levels 3 and 4) and the Programme for International Student Assessment (PISA) (i.e., levels 4 to 6) (Department of Education and Skills [DES], 2011).

Following the formulation of the 2011 Strategy, policy attention to high achievement in mathematics and science in Ireland has been ever-increasing. Additional policies have been established and other relevant documentation has emerged at a national level, such as *Ireland's National Skills Strategy 2025* (DES, 2016b), the *Action Plans for Education 2016–2019* and *2018* (Government of Ireland, 2018; DES, 2016a), the *Chief Inspector's 2018 report* (DES, 2018), and the *Report of the STEM Education Review Group* (The STEM Education Review Group, 2016). These reiterated the already established national targets for high achievement in mathematics and introduced more specific ones, while also setting corresponding targets for science.

After the interim review report of the 2011 National Strategy (DES, 2017a) and the development of the *STEM Education Policy Statement* in 2017 (DES, 2017b), high achievement in mathematics and science became a national priority in Ireland. National targets focusing on raising the proportions of students performing at the highest levels of proficiency and improving the scores of students performing at the highest national percentiles in mathematics and science in national and international large-scale assessments were introduced by the DES (DES, 2016b, 2017a, 2017b; Government of Ireland, 2018). Most educational policy documents on mathematics and/or science education across primary and post-primary levels in Ireland published since then have included these national targets.

While this policy attention to high achievement in mathematics and science might have been expected to prompt empirical research on the topic, research literature at a national level has been particularly limited, making the formulation of specific guidelines tailored to the needs of high achievers, in an effort to achieve the aforementioned targets, difficult. The few international research studies that examined high achievement in mathematics and/or science (see, for example, Kartal & Kutlu, 2017; Kourti, 2019; Tourón et al., 2018; Veas Iniesta et al., 2017), which could either provide some insights into the topic for Ireland or serve as a starting point for further Irish-based research, tended to focus on certain themes, such as sex differences (e.g., Ellison & Swanson, 2010; Stoet & Geary, 2013). It is also noteworthy that most of these studies have not involved in-depth analyses of data, such as examining a wide range of factors originating from different aspects of a student's life (e.g., individual, family, class, school) simultaneously or including domain-specific variables (e.g., mathematics anxiety, motivation for learning science) in the analysis. Also, very few studies examined the relationships of contextual variables with high achievement in a multiple regression context or applied robust statistical techniques (e.g., multilevel modelling) to account for the clustered nature of educational data and the particularities of large-scale datasets. Analyses that meet the aforementioned criteria could obtain more accurate and credible predictions of high achievement compared to other more descriptive analyses, while also mitigating potential limitations of the collected data (e.g., clustering).

Gilleece et al.'s (2010) study is the only one that set out to determine the role of a range of student- and school-level characteristics in high (and low) achievement in mathematics and science among Irish students. Gilleece et al.'s (2010) research made a significant contribution to the area; however, it was limited to one cohort of 15-year-old Irish students and employed a limited range of variables in the analysis. In acknowledging the limitations of their study and the need for further in-depth research on high achievement in mathematics and science in Ireland, Gilleece et al., (2010) provided a comprehensive list of recommendations for further research. Amongst others, the authors suggested that inclusion of contextual variables pertaining to individual domains (e.g., attitudes toward mathematics, career expectations related to science) in the analysis would provide a unique insight into the factors associated with high (and low) achievement in mathematics and science in Ireland. Although their recommendations had the potential to add further impetus to the investigation of high achievement in mathematics and science nationally, these recommendations have not yet been realised.

The current study

Prompted by the aforementioned policy attention and in an effort to examine the magnitude and consistency of issues related to high achievement in mathematics and science in Ireland, Pitsia et al., (2022) conducted an in-depth longitudinal investigation of high achievement across education levels and student cohorts using PISA, Trends in International Mathematics and Science Study (TIMSS), and Progress in International Reading Literacy Study (PIRLS) data for Ireland. The authors found that issues pertaining to high achievement in mathematics and science in Ireland have been more apparent and consistent at post-primary level, with issues at primary level being less clear-cut. Specifically, Pitsia et al., (2022) found that at post-primary level, proportions of high-achieving students in mathematics and science have not been in line with Ireland's good mean performance in the two subjects, having fewer high-achieving students compared to other similarly performing countries, with high achievers also often under-performing relative to their peers in other countries. In turn, these findings indicate that attention to high achievement at post-primary level should be considered a priority in Ireland.

Taking into account Gilleece et al.'s (2010) recommendations and Pitsia et al.'s (2022) findings, the study described in this paper is one of the first investigations of what may predict high achievement in mathematics and science among post-primary students in Ireland. The research question guiding the study is:

Which student, home, class, and school characteristics predict high achievement in mathematics and science among 15-year-old students in Ireland?

To answer this research question, the analysis of the data (i.e., model-building process) and the discussion of the results have been framed along the lines of the *socio-ecological model of human development* (Bronfenbrenner, 1994). According to this theoretical framework, students are nested within different environments with which they interact, and the components of these environments may, either directly or indirectly, contribute to their academic outcomes. Focusing on the *microsystem* of Bronfenbrenner's framework, which constitutes the most proximal environment experienced by students comprised of people and systems with which they have direct contact, the analysis of the data and the discussion of the results begin with variables stemming from the students themselves (e.g., demographics, self-beliefs, dispositions, drive, and engagement, mathematics and science-related activities) and their parents (e.g., parental dispositions and support), and they, next, focus on the variables stemming from the students' classes and schools (e.g., class disciplinary climate, school resources).

Within the context of ongoing efforts to introduce reforms in mathematics and science teaching and learning in Ireland, the results of this study are intended to extend the existing body of knowledge about high achievement, highlight the areas on which existing initiatives in mathematics and science education could focus to better address the needs of high achievers in these subjects, and prompt the formulation of new policies and initiatives that could assist towards this end.

Methods

Data

This paper presents results from a secondary analysis of PISA 2012 and 2015 data (Organisation for Economic Co-operation and Development [OECD], 2014, 2017b). PISA is a collaborative effort among the OECD countries to measure how well

15-year-old students are prepared to meet the challenges of the future (OECD, 2017b). The assessment, which is cross-sectional in nature, has taken place every three years since 2000, following a cyclical design within which it changes its major domain in every cycle, starting with reading in 2000, mathematics in 2003, and science in 2006. Ireland, which is the focus of this study, has participated in all PISA cycles since 2000. Data from the 2012 and 2015 cycles are used in this study as, at the time of writing, they are the most recent cycles in which mathematics and science were the major assessment domains, respectively.

Sampling and participants

PISA defines its international target population in terms of students' age, focusing on 15-year-old students in school in seventh grade or higher. These students are approaching the end of compulsory schooling in most participating countries and, thus, this age is considered as particularly important for personal and academic decisions, while school enrolment at this level is close to universal in most OECD countries (OECD, 2017b).

PISA selects its nationally representative samples based on a two-stage stratified sampling design. At the first stage, individual schools constitute the sampling units and are selected with probability proportional to size. The second-stage sampling units are students within sampled schools. Typically, 35 students or more are sampled with equal probability per school, although this size may vary depending on the total enrolment size of the school (OECD, 2017b). In Ireland, 5,015 15-year-olds in 182 schools took part in PISA 2012 and 5,741 15-year-olds in 167 schools took part in PISA 2015 (Perkins et al., 2013; Shiel et al., 2016); these constitute the samples involved in the analysis in the current study.

Measures and variables

PISA uses tests to assess, amongst others, 15-year-old students' reading, mathematics, and science achievement, providing an overall domain score and subdomain scores for each assessed student. The reading, mathematics, and science scales had been standardised to have a mean of 500 (on average across OECD countries) and a standard deviation (*SD*) of 100 when each was a major assessment domain for the first time (2000 for reading, 2003 for mathematics, and 2006 for science), with these figures slightly varying in subsequent PISA cycles. The concept of "literacy" is used in assessing and interpreting students' performance in the three domains. "Literacy" refers to "students' capacity to apply knowledge and skills in key subjects, and to analyse, reason and communicate effectively as they identify, interpret and solve problems in a variety of situations" (OECD, 2017a, p. 13).

In PISA, each student is administered a subset of the test items from the total item pool for each domain due to time restrictions. Consequently, different groups of students answer different, although overlapping, sets of items. Hence, student proficiencies are not observed; instead, they are missing data that must be inferred from the observed item responses. Given this design and to generate population-level proficiency estimates, PISA uses the imputation methodology of plausible values. Plausible values constitute random numbers drawn from the distribution of scores that could be reasonably assigned to each individual (Wu, 2005). Five plausible values were generated in PISA 2012 and ten plausible values were generated in PISA 2015.

In the current study, students' performance in each domain was treated as a discrete rather than a continuous outcome, with high achievers and non-high achievers as the two categories. This process was facilitated by the fact that PISA reports students' performance not only using continuous scores but also levels of performance to indicate the proficiency of students in a given domain. Apart from being allocated a specific score along the performance scales, groups of students are also allocated to specific proficiency levels and descriptions of the competencies associated with such levels are generally provided (see OECD, 2014, 2017b). In this study, mirroring OECD's approach, students scoring at proficiency levels 5 and 6 in each of the domains were identified as high achievers.

Prior to conducting any analysis, the main outcome variables had to be computed. In the PISA databases, binary variables indicating the proficiency level to which students belong do not exist; hence, using the relevant cut-off points for high achievement (i.e., proficiency level 5 in both mathematics and science) and in order to use all plausible value estimates of student performance in the analysis, each student was assigned the value 0 or 1 based on whether each plausible value estimate was below or above the established cut-off point for each domain, respectively. This was done separately for each plausible value estimate. These binary variables constituted the outcome variables of the analysis.

In addition to collecting information about students' academic literacy, PISA uses questionnaires to collect information about students' attitudes, interests, motivations, and beliefs as well as students' family and school backgrounds to facilitate interpretation of achievement outcomes (e.g., OECD, 2017b). In Ireland, student and school questionnaires were administered in 2012, and student, parent, teacher, and school questionnaires were administered in 2015. Information from these questionnaires was also used for the analysis in this study.

Statistical analysis

To examine the role of student- and school-level variables in students' high achievement in mathematics and science, PISA 2012 and 2015 data were analysed in two stages: (i) bivariate analysis and (ii) multilevel binary logistic regression analysis.

Bivariate analysis

In the first stage, bivariate analysis was performed to facilitate the decision-making about the inclusion of predictor variables into the next stage of the analysis (i.e., multilevel binary logistic regression analysis). The International Association for the Evaluation of Educational Achievement (IEA) International Database Analyzer (IDB Analyzer) (IEA, 2021) was used to calculate all estimates (e.g., percentages, means, standard errors [*SEs*], *SDs*) involved in the bivariate analysis. By using the IEA IDB Analyzer, the plausible values and replicate weights were used appropriately, and adjusted *SEs* were computed. The cut-off points for high achievement for each domain were used in the *benchmarks* statistic type in the IEA IDB Analyzer to facilitate treatment of student achievement as a binary outcome.

The effect size of each of the relationships between the predictor variables and the binary outcome variable in each subject (i.e., high achievement), rather than the statistical significance level alone, was the criterion for either including or not including variables in the next stage of the analysis. This was because with large sample sizes, such as the ones analysed in this study, very low coefficients are likely to be statistically significant. Guidelines provided by L. Cohen et al. (2017) and Fritz et al. (2012) related to the computation and reporting of effect sizes based on the types of variables and samples involved in the analysis were followed in the selection of the appropriate effect size measures. Based on these guidelines, the phi (ϕ) and the Cramer's V (ϕ_c) effect size measures were used for categorical variables with two or more categories, respectively; the eta-squared (η^2) effect size measure was used for ordinal variables; and the Hedges' g (g) effect size measure was used for the continuous variables due to the considerably different sample sizes of the two comparison groups (high and non-high achievers).

J. Cohen's (1988) in conjunction with Hattie's (2009) guidelines were used to identify the thresholds above which predictor variables would progress to the next stage of the analysis. The fact that effect sizes from different effect size families had to be used in this study meant that equivalent thresholds (i.e., thresholds that indicate effects of the same magnitude) were required. Predictor variables that yielded an effect size of 0.10 (ϕ and ϕ_c), 0.010 (η^2) or 0.20 (g) progressed to the multilevel binary logistic regression analysis as effect sizes at or above these thresholds are considered as *small, intermediate* or *large effects* (in general) and *teacher* or *desired effects* (in education).[1] Statistical significance tests and calculation of effect sizes for the estimates provided by the IEA IDB Analyzer were conducted using the relevant formulas for each estimate as recommended by the literature (see Gonzalez, 2014).

By selecting predictors for the multilevel binary logistic regression analysis based on both their statistical and practical significance, the aim was to build as parsimonious models as possible that examine the strongest predictors of high achievement in mathematics and science, and also signify the extent to which these predictors retain their significance after accounting for other variables.

Multilevel binary logistic regression analysis

Following the bivariate analysis, multilevel binary logistic regression analysis was conducted on Mplus (Muthén & Muthén, 2017) to examine the contribution of a range of factors stemming from the students, their parents, classes, and schools in the prediction of high achievement in mathematics and science. The sampling design that leads to the clustered nature of PISA samples means that students within the same schools may have more characteristics in common than with students from other schools (L. Cohen et al., 2017; OECD, 2017b). A statistic, the intra-class correlation (*ICC*), represents the proportion of the total variance in the outcome variable that is attributable to the cluster (here, school) (Field, 2018). Clustering constitutes a problem because many statistical models assume that cases are independent of each other, but students who study at the same school are less likely to be independent of each other. This clustering was taken into account in the regression analysis by conducting multilevel analysis, which can estimate the variation in the outcome variable that is attributable to differences within or between the clusters and identify the factors at each level that are associated with this influence, while not underestimating the *SE*s of the regression coefficients (L. Cohen et al., 2017;

[1] Student sex and socioeconomic status at the student and the school level were excluded from this criterion. These variables were used in the multilevel binary logistic regression analysis regardless of the effect sizes they yielded in the bivariate analysis.

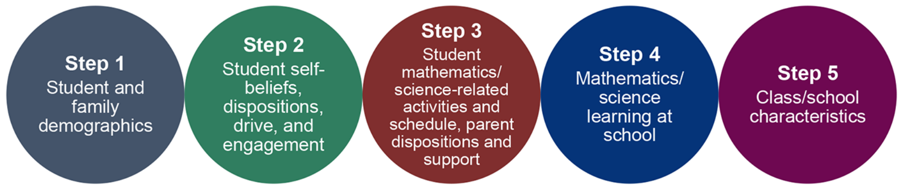

Fig. 1 Steps in building the hierarchical two-level binary logistic regression models

Woltman et al., 2012). Specifically, the contribution of student- and school-level variables in predicting post-primary students' high achievement in mathematics and science was evaluated through hierarchical two-level binary logistic regression models. Students were the unit of analysis at level 1 and schools were the unit of analysis at level 2.

The binary outcome variable used in this study can be conceptualised as the discretisation of an underlying continuous latent variable (i.e., high achievement/non-high achievement is a binary representation of the underlying continuous latent variable denoting the test score). Therefore, the continuous latent response variable approach for binary outcomes was used for the calculation of the estimates (McKelvey & Zavoina, 1975; Snijders & Bosker, 2012). Also, given that the continuous latent response variable approach for binary outcomes was employed, cluster sampling was performed in the assessments that were employed, and sampling weights were used in the analysis, parameters for the models were estimated using the maximum likelihood estimation with robust *SEs* and the logit link function (Muthén & Muthén, 2017).

Model-building strategy The first step taken in the model-building process was the construction of null models that contained no predictor variables at either level of the analysis for each cycle to compute the *ICCs* and, thus, determine the extent to which high achievement in mathematics and science can be attributed to between-student and between-school differences. Given that when modelling binary outcomes, the variability within groups (σ^2) is not normally distributed, the *ICCs* were estimated using the following approximation:

$$ICC = \frac{\tau_{00}}{\tau_{00} + (\pi^2/3)}$$

…where τ_{00} is the random intercept variance (i.e., the level-2 variance component) and $\pi^2/3(\approx 3.29)$refers to the variance of a standard logistic distribution (i.e., the assumed level-1 variance component; given that the logistic regression model does not include level-1 residual) (Goldstein, 2011; Goldstein et al., 2002; Hox et al., 2018; Sommet & Morselli, 2017).

The null models also served as a reference point against which the final models were compared, to evaluate the extent to which the addition of predictor variables contributed to the prediction of high achievement in mathematics and science.

Next, predictor variables were hierarchically entered into the models in blocks (see Fig. 1). Steps 1–4 involved student-level variables (level 1 of the analysis) and step 5 involved school-level variables (level 2 of the analysis). For the PISA 2012 data, there were four rather than five steps in the model as parent- and teacher-level variables were not measured in that cycle in Ireland.

Preliminary analysis (whereby variables were retained in all steps of the models regardless of their statistical significance) indicated that as the sample size decreased due to listwise deletion with each step of the model, the percentages of high achievers in the analysis samples increased. This might be linked to the fact that high achievers tend to have fewer missing responses compared to students with lower performance and, thus, more cases among the non-high achievers were being excluded from the analysis. Dropping the non-significant variables in every subsequent step of each model helped to mitigate this issue, resulting both into fewer missing cases and minor variations in the percentages of high achievers involved in the analysis (in the PISA 2012 model, percentages of high achievers across the steps ranged from 10.8 to 11.3% being very close to the percentage of high achievers in the overall sample [10.7%]; similarly, in the PISA 2015 model, percentages of high achievers across the steps ranged from 7.2 to 7.7% being very close to the percentage of high achievers in the overall sample [7.1%]).

The final iteration of the modelling process explored the possibility of statistically significant interactions between predictor variables in predicting the outcome. Due to the lack of existing research on the significance of certain interactions in the prediction of high achievement in mathematics and science, this process was exploratory in nature and examined interactions that could be meaningful for policy and practice purposes. Thus, the significance of cross-level interactions of student sex and student-level socioeconomic status with the school-level variables included in the analysis was explored. The interactions terms were entered one by one into each final model regardless of whether the main effects were statistically significant or not because, as L. Cohen et al., (2017) argue, an interaction effect may occur even when no main effects are present.[2]

Weights, plausible values, and centring

As recommended by the relevant research literature, sampling and replicate weights (where appropriate) were used in the analysis. The appropriate use of weights for the bivariate analysis was secured through generating all necessary estimates using the IEA IDB Analyzer, while in the multilevel binary logistic regression analysis, Asparouhov's (2009) recommendations for the use of sampling weights at the two levels of the analysis were followed. Specifically, the final student weight was decomposed into a "within-school weight" and a "between-school weight" for the appropriate estimation. This was done by dividing the final student weight by the final school weight. The resulting "within-school weight" was used as the level-1 weight and the original final school weight was used as the level-2 weight (see also *Withincluster Weights* method in Mang et al., 2021).

The use of plausible values in the analysis adhered to the relevant procedures outlined in the PISA technical reports and data analysis manual, but also in von Davier et al.'s (2009) study. Specifically, all plausible values were used for the estimation of student achievement, and the imputation variance was taken into account in the estimation of the *SEs*. The correct use of the plausible values for the bivariate analysis was secured through generating all necessary estimates using the IEA IDB Analyzer. For the multilevel modelling, imputation techniques were applied to involve all plausible values in the analysis (i.e., *type*=*imputation*).

[2] In the exploration of the statistical significance of interaction terms, corrections for multiple comparisons were applied (Armstrong, 2014).

Continuous predictor variables were centred for the multilevel modelling. Given that the goal of the analysis was to examine the absolute (between-student) contribution of variables in the prediction of high achievement in mathematics and science, grand-mean centring was used (Sommet & Morselli, 2017).

Results

Bivariate analysis

Bivariate analysis was conducted to examine the relationships of a range of contextual variables stemming from students, their parents, teachers, and schools with students' high achievement in mathematics and science and facilitate the decision-making about the inclusion of predictor variables into the multilevel binary logistic regression analysis. All contextual variables included in the PISA 2012 and 2015 databases with data for Ireland were included in the bivariate analysis. Background variables that (i) had a statistically significant relationship with the outcome variable and (ii) yielded effect sizes of 0.10 (ϕ and ϕ_c), 0.010 (η^2) or 0.20 (g) progressed to the multilevel binary logistic regression analysis. Tables 1, 2, 3 and 4 present the percentages of students in each performance group (high achievers and non-high achievers) and the means and *SDs* of high achievers and non-high achievers for each of the background variables that met these criteria and, thus, were included in the multilevel binary logistic regression analysis along with student sex and socioeconomic status (which were always included in the models) at the student and school levels. The tables also include the results of the tests that were conducted to examine the statistical significance of the differences in the distribution of students across each background variable between the two performance groups or the statistical significance of the differences in the means for each of the background variables between high achievers and non-high achievers.[3]

In PISA 2012, none of the categorical variables examined yielded large effect sizes. In PISA 2015, however, there was a number of categorical background variables in which the distribution of high and non-high achievers across the different variable categories differed (Tables 1 and 2). Specifically, a set of variables related to the students' family involvement in science-related careers and their parents' perceptions and expectations in relation to science were found to be significantly related to students' belonging to either the high or the non-high-achieving science performance group (Table 1). Although parents' perceptions and expectations yielded the largest differences between the two performance groups among the examined variables, the magnitude of these differences was relatively small (ϕ=0.18 and ϕ=0.19). In general, there were approximately four times more high achievers in the groups of students whose parents had more positive perceptions of and higher expectations about science compared to the rest of the students. Additionally, there were significantly fewer high achievers in science among the students who reported watching TV/DVD/video, using the internet, chats or social networks, meeting or talking to friends on the phone, and exercising or practising a sport on school days, either before or after going to school.

Among the ordinal variables examined, student educational expectations in the 2015 cycle was the only one that reached the effect size threshold of η^2=0.010 (Table 2), with

[3] Although these results provide insights into the relationships of a range of contextual variables with high achievement in mathematics and science, they do not provide information about the direction of these relationships, which means that any observed relationships may be reciprocal.

Table 1 Distribution of students across nominal background variables by performance group

			High achievers		Non-high achievers			
Variable	**Categories**	**N**	**%**	***n*observed (expected)**	**%**	***n*observed (expected)**	**χ^2**	**φ**
PISA 2012 - mathematics								
Student sex	females	2,471	8.9	220 (263)	91.1	2,251 (2,208)	15.19	0.06
	males	2,545	12.3	314 (271)	87.7	2,231 (2,274)		
PISA 2015 - science								
Parent perceptions and expectations: do you expect your child will go into a science-related career?	no	3,051	3.5	108 (230)	96.5	2,943 (2,821)	180.17	0.19
	yes	1,928	13.9	268 (146)	86.1	1,660 (1,782)		
Parent perceptions and expectations: does your child show interest in working in a science-related career?	no	2,709	2.9	79 (202)	97.1	2,630 (2,507)	174.38	0.19
	yes	2,335	12.8	298 (175)	87.2	2,037 (2,160)		
Parent perceptions and expectations: has your child shown interest in studying science after completing secondary school?	no	2,804	3.2	89 (211)	96.8	2,715 (2,593)	170.63	0.18
	yes	2,224	13.0	289 (167)	87.0	1,935 (2,057)		
Parent perceptions and expectations: do you expect your child will study science after completing secondary school?	no	2,893	3.5	100 (217)	96.5	2,793 (2,676)	160.15	0.18
	yes	2,091	13.1	273 (156)	86.9	1,818 (1,935)		
Before going to school: Internet\Chat\Social networks (e.g. Facebook)	no	1,780	11.6	206 (129)	88.4	1,574 (1,651)	72.32	0.11
	yes	3,742	5.2	194 (271)	94.8	3,548 (3,471)		
Member of student's family (including parent) working in a science-related career	no	3,318	5.4	180 (250)	94.6	3,138 (3,068)	60.64	0.11
	yes	1,740	11.6	201 (131)	88.4	1,539 (1,609)		
After leaving school: Meet friends or talk to friends on the phone	no	1,195	12.3	147 (86)	87.7	1,048 (1,109)	58.04	0.10
	yes	4,278	5.8	248 (309)	94.2	4,030 (3,969)		
Before going to school: Exercise or practise a sport	no	3,571	9.2	330 (260)	90.8	3,241 (3,311)	57.55	0.10
	yes	1,892	3.6	68 (138)	96.4	1,824 (1,754)		
Before going to school: Meet friends or talk to friends on the phone	no	2,844	9.8	279 (206)	90.2	2,565 (2,638)	57.24	0.10
	yes	2,638	4.5	118 (191)	95.5	2,520 (2,447)		
Before going to school: Watch TV\DVD\Video	no	3,679	9.1	333 (267)	90.9	3,346 (3,412)	52.05	0.10
	yes	1,814	3.6	66 (132)	96.4	1,748 (1,682)		
Student sex	females	2,833	4.9	139 (200)	95.1	2,694 (2,633)	39.26	0.08
	males	2,908	9.2	267 (206)	90.8	2,641 (2,702)		

Notes. Variables in descending order of effect size. All differences were statistically significant at the 0.01 level

high achievers in science having significantly higher educational expectations compared to their non-high-achieving peers.

Results presented in Tables 3 and 4 reveal that the patterns for mathematics and science were highly consistent across the two PISA cycles with regards to differences between high and non-high achievers in the continuous background variables. Student self-beliefs and attitudes, including student mathematics and science self-efficacy, self-concept, anxiety, and enjoyment, their environmental awareness, openness for problem-solving, familiarity with mathematical concepts, epistemological beliefs about science, and interest in broad science topics were among those with the highest effect sizes. High achievers had higher self-efficacy, self-concept, enjoyment, environmental awareness,

Table 2 Distribution of students across ordinal background variables

			High achievers			Non-high achievers					
Variable	**Categories**	**N**	***%***	***n***	***mean rank***	***%***	***n***	***mean rank***	***U***	***Z***	***η²***
PISA 2015 - science											
Student educational expectations	Lower secondary education	721	1.9	14	3875.0	98.1	707	2756.4	1490634.5	-13.26	0.031
	Leaving Cert Applied, Transition year, VTOS and FÁS programmes	258	1.6	4		98.5	254				
	Leaving Certificate and Vocational programmes	791	1.6	13		98.4	778				
	Post-secondary, non-tertiary	219	1.4	3		98.6	216				
	Tertiary (National Framework of Qualifications [NFQ] levels 6 (higher) and 7)	1,070	4.2	45		95.8	1,025				
	Tertiary (NFQ level 8)	2,613	12.5	327		87.5	2,286				

Notes. Difference was statistically significant at the 0.01 level

openness for problem-solving, familiarity with mathematical concepts, epistemological beliefs about science, and interest in broad science topics and lower levels of mathematics anxiety compared to non-high achievers. The magnitude of these differences ranged from moderate to very large. Effect sizes of these differences between the two performance groups were relatively larger for mathematics compared to science, indicating that the two performance groups in mathematics differed in these non-cognitive factors to a greater extent compared to the two performance groups in science.

Students' family socioeconomic background was also significantly different for the two performance groups in both subjects (PISA 2012: $g=0.70$; PISA 2015: $g=0.74$), with high achievers coming from families of more affluent socioeconomic background compared to their non-high-achieving peers.

Results also indicated that high achievement in mathematics and science is not only dependent on individual characteristics but on class and school characteristics too. Class and school characteristics such as disciplinary climate, school autonomy, school size, teaching time, and school socioeconomic composition were significantly different between the two performance groups in mathematics and science, suggesting that students in the two performance groups tended to attend somewhat different classes and schools.

Table 3 Means in continuous background variables by mathematics performance group, PISA 2012

Variable		N	*M (SD)*	*MD*	*SED*	*95% CI*	*t*	*g*
Student mathematics self-efficacy	high achievers	346	1.05 (0.83)	1.16	0.05	1.06, 1.26	22.64	1.29
	non-high achievers	2,958	-0.11 (0.91)					
Student mathematics self-concept	high achievers	365	0.86 (0.90)	1.01	0.05	0.91, 1.11	20.64	1.14
	non-high achievers	2,962	-0.15 (0.88)					
Student mathematics anxiety	high achievers	365	-0.72 (0.98)	-0.93	0.05	-1.02, -0.84	19.38	1.07
	non-high achievers	2,960	0.21 (0.85)					
Student openness for problem solving	high achievers	347	0.82 (0.87)	0.94	0.05	0.84, 1.04	17.93	1.02
	non-high achievers	2,947	-0.12 (0.93)					
Student familiarity with mathematical concepts	high achievers	352	0.21 (0.74)	0.75	0.05	0.65, 0.85	14.59	0.82
	non-high achievers	2,957	-0.54 (0.93)					
Family economic, social and cultural status	high achievers	533	0.64 (0.72)	0.58	0.04	0.51, 0.65	15.28	0.70
	non-high achievers	4,440	0.06 (0.84)					
Student interest in mathematics	high achievers	346	0.62 (0.91)	0.63	0.05	0.52, 0.74	11.72	0.67
	non-high achievers	2,960	-0.01 (0.95)					
Student perseverance	high achievers	347	0.61 (1.01)	0.52	0.06	0.41, 0.63	9.07	0.51
	non-high achievers	2,951	0.09 (1.01)					
Student instrumental motivation in mathematics	high achievers	346	0.50 (0.88)	0.42	0.05	0.32, 0.52	8.07	0.46
	non-high achievers	2,957	0.08 (0.92)					
Student mathematics work ethic	high achievers	346	0.45 (0.90)	0.43	0.05	0.33, 0.53	8.16	0.46
	non-high achievers	2,955	0.02 (0.93)					
Student attributions to failure in mathematics	high achievers	347	-0.46 (0.87)	-0.40	0.05	-0.51, -0.29	7.41	0.42
	non-high achievers	2,953	-0.06 (0.96)					
Student mathematics behaviour	high achievers	347	-0.08 (0.82)	0.39	0.06	0.28, 0.50	7.06	0.40
	non-high achievers	2,957	-0.47 (0.99)					
Disciplinary climate in mathematics classes	high achievers	366	0.46 (1.02)	0.37	0.06	0.25, 0.49	6.12	0.34
	non-high achievers	2,965	0.09 (1.10)					
Student mathematics intentions	high achievers	343	0.12 (0.90)	0.26	0.06	0.15, 0.37	4.73	0.27
	non-high achievers	2,889	-0.14 (0.97)					
Student-related factors affecting school climate	high achievers	502	0.12 (0.91)	0.23	0.04	0.15, 0.31	5.34	0.25
	non-high achievers	4,092	-0.11 (0.91)					
School size	high achievers	534	662.54 (260.83)	58.40	12.37	34.11, 82.69	4.72	0.22
	non-high achievers	4,482	604.14 (271.18)					
ICT availability at school	high achievers	531	-0.22 (0.73)	-0.17	0.04	-0.25, -0.09	4.37	0.20
	non-high achievers	4,404	-0.05 (0.86)					
Shortage of educational staff in the school	high achievers	502	-0.30 (0.82)	-0.17	0.04	-0.25, -0.09	4.29	0.20
	non-high achievers	4,092	-0.13 (0.84)					
School mean of family economic, social and cultural status	high achievers	534	-0.59 (0.20)	0.00	0.01	-0.02, 0.02	0.00	0.00
	non-high achievers	4,482	-0.59 (0.20)					

Notes. Variables in descending order of effect size. All differences were statistically significant at the 0.01 level except for the difference in the *school mean of family economic, social and cultural status* variable for which the difference was not statistically significant. *MD*: mean difference, *SED*: SE of mean difference

Multilevel binary logistic regression analysis

Following the bivariate analysis, hierarchical two-level binary logistic regression models were applied to the PISA 2012 and 2015 data. All variables that reached the effect size thresholds for each PISA cycle were included in the models.

Given the sampling procedures followed by PISA, whereby students are nested within the same schools (i.e., clusters), variation in achievement (and hence, high achievement) in these assessments can be separated into between-student and between-school components. The *ICCs*, which indicate the extent to which schools differ with respect to

Table 4 Means in continuous background variables by science performance group, PISA 2015

Variable		**N**	*M (SD)*	*MD*	*SED*	*95% CI*	*t*	*g*
Student enjoyment of science	high achievers	404	1.04 (0.90)	0.91	0.06	0.80, 1.02	16.35	0.84
	non-high achievers	5,159	0.13 (1.09)					
Student science self-efficacy	high achievers	402	0.92 (0.95)	0.93	0.06	0.81, 1.05	15.29	0.79
	non-high achievers	5,108	-0.01 (1.19)					
Epistemological beliefs about science	high achievers	402	0.79 (0.83)	0.63	0.04	0.54, 0.72	14.49	0.75
	non-high achievers	5,111	0.16 (0.84)					
Student interest in broad science topics	high achievers	402	0.70 (0.68)	0.69	0.05	0.60, 0.78	14.42	0.75
	non-high achievers	5,093	0.01 (0.94)					
Family economic, social and cultural status	high achievers	404	0.73 (0.72)	0.61	0.04	0.53, 0.69	14.36	0.74
	non-high achievers	5,263	0.12 (0.83)					
Student science activities	high achievers	401	0.35 (0.87)	0.78	0.06	0.67, 0.89	14.24	0.74
	non-high achievers	5,146	-0.43 (1.07)					
Students' past science activities	high achievers	383	0.58 (0.84)	0.63	0.05	0.53, 0.73	12.10	0.64
	non-high achievers	4,709	-0.05 (0.99)					
Students' expected occupational status	high achievers	356	68.10 (13.19)	9.81	0.89	8.06, 11.56	11.03	0.61
	non-high achievers	4,618	58.29 (16.37)					
School mean of family economic, social and cultural status	high achievers	406	0.35 (0.39)	0.21	0.02	0.17, 0.25	10.98	0.57
	non-high achievers	5,335	0.14 (0.37)					
Parent view on science	high achievers	382	0.93 (0.98)	0.60	0.06	0.48, 0.72	10.16	0.54
	non-high achievers	4,683	0.33 (1.12)					
Student test anxiety	high achievers	405	-0.28 (0.87)	-0.46	0.05	-0.55, -0.37	10.15	0.52
	non-high achievers	5,272	0.18 (0.88)					
Student achievement motivation	high achievers	405	0.80 (0.93)	0.44	0.05	0.35, 0.53	9.46	0.49
	non-high achievers	5,269	0.36 (0.90)					
Student instrumental motivation in science	high achievers	403	0.79 (0.92)	0.46	0.05	0.36, 0.56	9.11	0.47
	non-high achievers	5,118	0.33 (0.98)					
Student environmental awareness	high achievers	405	0.77 (1.02)	0.49	0.06	0.38, 0.60	8.46	0.44
	non-high achievers	5,190	0.28 (1.13)					
Student perceived autonomy related to ICT use	high achievers	398	0.48 (0.92)	0.40	0.05	0.31, 0.49	8.35	0.43
	non-high achievers	5,033	0.08 (0.92)					
Student value of co-operation	high achievers	404	-0.35 (0.98)	-0.42	0.05	-0.52, -0.32	8.15	0.42
	non-high achievers	5,262	0.07 (1.00)					
Percentage of students from socioeconomically disadvantaged homes in school	high achievers	356	20.54 (19.03)	-8.93	1.26	-11.40, -6.46	7.10	0.39
	non-high achievers	4,735	29.47 (23.14)					
Teacher fairness	high achievers	405	8.75 (3.03)	-1.12	0.19	-1.50, -0.74	5.82	0.30
	non-high achievers	5,268	9.87 (3.78)					
Parental current support for learning at home	high achievers	383	0.05 (0.73)	0.23	0.04	0.14, 0.32	5.20	0.28
	non-high achievers	4,727	-0.18 (0.84)					
Average time per week on science	high achievers	404	161.18 (76.00)	19.01	3.80	11.55, 26.47	5.01	0.26
	non-high achievers	5,221	142.17 (73.29)					
Student-related factors affecting school climate	high achievers	382	-0.17 (0.99)	-0.23	0.05	-0.32, -0.14	4.78	0.25
	non-high achievers	4,946	0.06 (0.90)					
Student perceived ICT competence	high achievers	399	0.41 (0.92)	0.21	0.05	0.12, 0.30	4.48	0.23
	non-high achievers	5,056	0.20 (0.90)					
School autonomy	high achievers	388	0.78 (0.14)	0.03	0.01	0.02, 0.04	4.35	0.23
	non-high achievers	5,015	0.75 (0.13)					
ICT availability at school	high achievers	385	5.48 (1.81)	-0.47	0.11	-0.69, -0.25	4.21	0.22
	non-high achievers	4,736	5.95 (2.13)					

Table 4 (continued)

Variable		N	*M (SD)*	*MD*	*SED*	*95% CI*	*t*	*g*
Shortage of educational material	high achievers	387	0.01 (1.10)	-0.26	0.06	-0.38, -0.14	4.13	0.22
	non-high achievers	4,978	0.27 (1.20)					
Adaption of instruction	high achievers	395	0.16 (0.88)	0.20	0.05	0.10, 0.30	4.04	0.21
	non-high achievers	4,629	-0.04 (0.95)					
Availability of computers at school	high achievers	382	0.58 (0.30)	-0.09	0.02	-0.14, -0.04	3.84	0.20
	non-high achievers	4,884	0.67 (0.45)					
Science specific resources	high achievers	391	5.98 (1.44)	0.32	0.08	0.16, 0.48	3.81	0.20
	non-high achievers	4,996	5.66 (1.61)					

Notes. Variables in descending order of effect size. All differences were statistically significant at the 0.01 level. *MD*: mean difference, *SED*: SE of mean difference

high achievement, were calculated through a *null model*, which contained no predictors at either level of the analysis, for each cycle. In 2012 and 2015, 12.2% and 13.8% of the variance in student high achievement in mathematics and science, respectively, was attributed to between-school differences. The significant proportion of variance in high achievement in mathematics and science attributed to the cluster warranted the consideration of the hierarchical nature of the data through conducting multilevel regression models.

The tables that summarise the results of the hierarchical two-level binary logistic regression models below present the proportions of variance (R^2; expressed as a percentage of the total variance)[4] in high achievement explained at each level by each step of the models, the *threshold* (statistic equivalent to the more commonly used *intercept*, with the two being the same except that they have opposite signs; Muthén & Muthén, 2017), the standardised coefficients (*βs*) accompanied by their *SEs* for each predictor variable, the odds ratios (*ORs*) for the statistically significant variables in the final model, and the fit statistics for each step of the models, including the null models. Although all steps of the models are presented in the tables, only results from the final models and *ORs* based on the coefficients of the final models are discussed below.[5]

PISA 2012 – mathematics

Table 5 summarises the results of the hierarchical two-level binary logistic regression model for high achievement in mathematics that was applied to PISA 2012 data. It should be noted here that in PISA 2012, rotated student questionnaires covering attitudinal and other non-cognitive constructs were used to increase the content coverage without increasing the response time for individual students (OECD, 2014). Due to this rotated design, there was an increased number of missing cases for these variables (i.e., missing data by design). Preliminary exploratory analysis, whereby attitudinal variables were entered into the model individually, revealed that the missing patterns for the *student familiarity with mathematical concepts* variable were different from those for other variables, which precluded the coexistence of this variable with the other variables

[4] R^2 at both levels of the analysis was computed based on McKelvey and Zavoina's (1975) and Snijders and Bosker's (2012) instructions for the estimation of R^2 in multilevel logistic regression models whereby the binary outcome variable is generated through the dichotomisation of an underlying continuous latent variable. Mplus calculates R^2 in line with these instructions (Muthén & Muthén, 2017). With the caveat that the R^2 refers to this continuous latent outcome variable as opposed to the observed binary outcome, its interpretation is similar to the interpretation of the R^2 for linear regression. For further technical information, see McKelvey and Zavoina (1975) and Snijders and Bosker (2012).

[5] Across both models, none of the cross-level interactions of student sex and student-level socioeconomic status with the school-level variables that were examined reached statistical significance.

Table 5 Hierarchical two-level binary logistic regression model for high achievement in mathematics, PISA 2012

		Step 1	Step 2	Step 3	Step 4	
R^2	student-level (%)	15.9	46.4	45.0	41.8	
	school-level (%)				36.4	
Threshold (*SE*)		2.31 (0.09)	3.28 (0.29)	3.32 (0.23)	3.25 (0.22)	
Student-level variables (reference category)		***β*(*SE*)**	***β* (*SE*)**	***β* (*SE*)**	***β* (*SE*)**	***OR***
Student sex (male)		−0.24 (0.08)**	−0.04 (0.11)			
Family economic, social and cultural status		0.38 (0.04)***	0.23 (0.05)***	0.25 (0.05)***	0.22 (0.05)***	1.87
Student mathematics anxiety			−0.25 (0.07)***	−0.22 (0.05)***	−0.23 (0.05)***	0.53
Student attributions to failure in mathematics			0.02 (0.05)			
Student instrumental motivation in mathematics			0.01 (0.06)			
Student interest in mathematics			−0.12 (0.07)			
Student mathematics behaviour			−0.04 (0.07)			
Student mathematics self-efficacy			0.30 (0.06)***	0.30 (0.06)***	0.32 (0.06)***	2.22
Student mathematics intentions			−0.01 (0.06)			
Student mathematics work ethic			−0.10 (0.07)			
Student openness for problem-solving			0.21 (0.07)**	0.17 (0.06)**	0.17 (0.06)**	1.55
Student perseverance			−0.02 (0.05)			
Student mathematics self-concept			0.14 (0.09)			
Disciplinary climate in mathematics classes				0.08 (0.06)		
ICT availability at school				−0.08 (0.05)		
School-level variables						
School mean of family economic, social and cultural status					−0.16 (0.26)	
School size					0.27 (0.26)	
Student-related factors affecting school climate					0.37 (0.42)	
Shortage of educational staff in the school					−0.09 (0.26)	
Fit statistics	Loglikelihood (H_0)	-1455.30	-349.19	-359.24	-337.36	
	AIC	2918.59	728.39	734.49	694.72	
	BIC	2944.64	809.00	777.64	747.83	

Note. Null model: *Threshold* (*SE*): 2.45 (0.11), H_0 = -1557.69, *AIC* = 3119.38, *BIC* = 3132.42. *$p<.05$; **$p<.01$; ***$p<.001$

into the model. Therefore, a decision was made to exclude this variable from the analysis. Despite this exclusion, there was still relatively large missingness in the PISA 2012 model, due to the rotated design. However, the percentage of high achievers in the sample involved in the analysis varied from 10.8 to 11.3% across the different steps of the model, being very close to the percentage of high achievers in the overall sample (10.7%), thus, confirming that the analysis sample was representative of the overall sample despite the relatively large missingness.

After accounting for the other predictors, four student-level variables were significantly associated with the odds of students being high as opposed to non-high achievers in mathematics in the final model. One of the variables related to students' socioeconomic background and the remaining three variables related to students' self-beliefs and engagement. The *ORs* in the final model (step 4) indicated that with an increase of one unit (i.e., one point) in the students' *family economic, social, and cultural status* index, students were 87% more likely to belong to the high-achieving compared to the non-high-achieving group in mathematics ($OR = 1.87$). In other words, students who were otherwise identical with regards to all the other variables included in the model but had a higher economic, social, and cultural status were much more likely to be high achievers in mathematics. Additionally, with an increase of one unit in the *student mathematics anxiety* index, students who were otherwise identical with regards to all the other variables were 47% less likely to be high achievers in mathematics, while students were 2.2 times more likely to be high achievers in mathematics, with every extra unit increase in the *self-efficacy* index. Finally, every one-unit increase in the *student openness for problem-solving* variable was associated with a 55% greater chance of high achievement in mathematics among 15-year-olds.

Student sex was a statistically significant predictor of 15-year-olds' high achievement in mathematics, with sex differences favouring boys, when first entered into the model (step 1), and after accounting for students' family socioeconomic status. However, sex differences were no longer significant when students' self-beliefs, dispositions, drive, and engagement were taken into account.

At school-level, none of the four variables included in the model significantly predicted students' odds of belonging to the high-achieving group in mathematics in PISA 2012. This indicated that the proportion of school-level variance in students' high achievement in mathematics that was explained by the model (36.4%) might have, in fact, been explained by the student-level variables that were already included in the model rather than the school-level variables.

The final model accounted for 41.8% of the between-student (level-1) and 36.4% of the between-school (level-2) variance in high achievement in mathematics, indicating that the student- and school-level variables that were included in the model contributed considerably to the prediction of the odds of 15-year-old students belonging to the high-achieving as opposed to the non-high-achieving group in mathematics.

PISA 2015 – science

Table 6 summarises the results of the hierarchical two-level binary logistic regression model for high achievement in science that was applied to PISA 2015 data. In the final model, a range of student-level characteristics were significantly associated with the odds of 15-year-old students belonging to the high-achieving group in science, while none of the variables at the school level retained their statistical significance, after partialling out the variability in the outcome variable due to the other variables included in the model. After accounting for other predictors, students' educational expectations, their interest in broad science topics, and their engagement in exercising or playing a sport before going to school were the strongest predictors of high achievement in science, yielding the largest βs at the student level.

Specifically, students who reported that they expected to complete *tertiary education at NFQ level 8* (i.e., higher diploma, honours bachelor's degree) were 2.9[6] and 3.46 times more likely to belong to the high-achieving group in science compared to students who expected to complete up to *lower secondary education* and *Leaving Certificate and Vocational programmes* (i.e., upper secondary education), respectively. Additionally, students'

[6] Calculated based on the inverted OR (1/OR).

Table 6 Hierarchical two-level binary logistic regression model for high achievement in science, PISA 2015

		Step 1	**Step 2**	**Step 3**	**Step 4**	**Step 5**	
R^2	student-level (%)	19.8	48.2	52.0	51.3	50.6	
	school-level (%)					83.6	
Threshold (*SE*)		2.77 (0.14)	3.05 (0.22)	2.70 (0.34)	2.63 (0.26)	2.73 (0.26)	
Student-level variables (reference category)		***β* (*SE*)**	***β* (*SE*)**	***β* (*SE*)**	***β* (*SE*)**	***β*(*SE*)**	***OR***
Student sex (male)		−0.31 (0.07)***	−0.18 (0.08)*	−0.17 (0.09)*	−0.19 (0.08)*	−0.22 (0.08)**	0.57
Family economic, social and cultural status		0.42 (0.04)***	0.18 (0.04)***	0.19 (0.04)***	0.20 (0.04)***	0.15 (0.04)**	1.57
Student environmental awareness			−0.12 (0.05)**	−0.11 (0.05)*	−0.11 (0.05)*	−0.10 (0.05)*	0.79
Student enjoyment of science			0.14 (0.05)**	0.08 (0.05)			
Student interest in broad science topics			0.17 (0.05)**	0.13 (0.06)*	0.20 (0.04)***	0.23 (0.04)***	1.90
Student instrumental motivation in science			0.04 (0.04)				
Student science self-efficacy			0.12 (0.05)*	0.12 (0.05)*	0.15 (0.05)**	0.14 (0.05)**	1.36
Epistemological beliefs about science			0.14 (0.04)***	0.14 (0.04)***	0.14 (0.04)***	0.14 (0.04)***	1.53
Students' expected occupational status			0.08 (0.04)				
Student test anxiety			−0.15 (0.04)***	−0.13 (0.04)**	−0.12 (0.04)**	−0.13 (0.04)**	0.69
Student achievement motivation			0.04 (0.04)				
Student value of co-operation			−0.15 (0.04)***	−0.11 (0.04)**	−0.13 (0.04)**	−0.11 (0.04)**	0.75
Student perceived ICT competence			−0.08 (0.05)				
Student perceived autonomy related to ICT use			0.08 (0.04)				
Student educational expectations (Tertiary (NFQ level 8))							
Lower secondary education			−0.36 (0.17)*	−0.37 (0.18)*	−0.37 (0.17)*	−0.42 (0.17)*	0.34
Leaving Certificate Applied, Transition Year, VTOS and FÁS programmes			−0.38 (0.40)	−0.36 (0.33)	−0.31 (0.31)	−0.28 (0.33)	
Leaving Certificate and Vocational programmes			−0.48 (0.22)*	−0.38 (0.17)*	−0.49 (0.21)*	−0.48 (0.20)*	0.29
Post-secondary, non-tertiary			−0.50 (0.37)	−0.41 (0.29)	−0.37 (0.29)	−0.28 (0.30)	
Tertiary - NFQ levels 6 (higher) and 7			−0.21 (0.12)	−0.20 (0.11)	−0.24 (0.11)	−0.23 (0.11)	
Before going to school: Watch TV\DVD\Video (no)				−0.16 (0.10)			
Before going to school: Internet\Chat\Social networks (e.g. Facebook) (no)				−0.08 (0.07)			
Before going to school: Meet friends or talk to friends on the phone (no)				−0.06 (0.09)			
Before going to school: Exercise or play a sport (no)				−0.28 (0.10)**	−0.36 (0.10)***	−0.36 (0.10)***	0.40

Table 6 (continued)

		Step 1	Step 2	Step 3	Step 4	Step 5	
After leaving school: Meet friends or talk to friends on the phone (no)				−0.09 (0.07)			
Member of student's family (including parent) working in a science-related career (no)				−0.01 (0.08)			
Student-level variables (reference category)		***β*(*SE*)**	***β*(*SE*)**	***β*(*SE*)**	***β*(*SE*)**	***β*(*SE*)**	***OR***
Parent perceptions and expectations: child shows interest in working in a science-related career (no)				0.12 (0.15)			
Parent perceptions and expectations: expect child will go into a science-related career (no)				0.16 (0.16)			
Parent perceptions and expectations: child shows interest in studying science after completing secondary school (no)				0.01 (0.17)			
Parent perceptions and expectations: expect child will study science after completing secondary school (no)				−0.03 (0.19)			
Student science activities				0.01 (0.05)			
Students' past science activities				0.13 (0.05)**	0.14 (0.05)*	0.15 (0.05)**	1.47
Parental current support for learning at home				−0.05 (0.04)			
Parent view on science				0.01 (0.05)			
Average time per week on science					−0.01 (0.04)		
Adaption of instruction					0.03 (0.04)		
ICT availability at school					−0.12 (0.04)**	−0.04 (0.02)*	0.90
Teacher fairness					−0.13 (0.05)**	−0.04 (0.01)**	0.90
School-level variables							
School mean of family economic, social and cultural status						0.49 (0.27)	
Availability of computers at school						−0.44 (0.33)	
School autonomy						−0.05 (0.22)	
Shortage of educational material						−0.33 (0.23)	
Science specific resources						0.03 (0.25)	
Student-related factors affecting school climate						−0.09 (0.22)	
Fit statistics	Loglikelihood (H_0)	-1236.91	-855.42	-800.96	-788.21	-739.89	
	AIC	2481.82	1752.84	1661.93	1618.42	1529.78	
	BIC	2508.39	1887.54	1853.19	1749.99	1686.10	

Notes. The variable *percentage of students from socioeconomically disadvantaged homes in school* was excluded from the model due to multicollinearity with the *school mean of family economic, social and cultural status* variable. Null model: *Threshold* (*SE*): 2.92 (0.13), H_0 = -1342.21, *AIC* = 2688.41, *BIC* = 2701.72. *$p<.05$; **$p<.01$; ***$p<.001$

odds of being high achievers would be increased by 90% with every extra unit in the *interest in broad science topics* index. Finally, students who reported exercising before going to school were 60% less likely to be high achievers in science compared to their peers who were otherwise identical with regards to the other variables but did not exercise or play a sport before school.

Student sex along with a number of other student-level non-cognitive and attitudinal factors also remained significant predictors of high achievement in science in the final model. Females were 43% less likely to be high achievers in science compared to their male peers. Also, greater environmental awareness, higher test anxiety, and value of co-operation were linked to decreased odds of a 15-year-old student belonging to the high-achieving group in science. Higher science self-efficacy and more positive views on scientific approaches, on the other hand, were associated with increased odds of 15-year-old students being high achievers in science. Finally, more frequent engagement of students in science-related learning activities at home at the age of 10, according to their parents' reports, was linked to increased odds of belonging to the high-achieving group in science; students were 1.5 times more likely to be high achievers in science, with every extra unit in the *past science activities* index.

Even though student enjoyment of science was the variable that yielded the strongest effect size ($g=0.84$) in the bivariate analysis for PISA 2015, and was a significant predictor of student high achievement in science when initially entered into the model (step 2), it was no longer significant after the introduction of additional variables in step 3. This indicates that variables such as students' activities and schedule on school days as well as science activities at a younger age, parental expectations and support at home may have accounted for the contribution of student enjoyment of science to the prediction of high achievement in the subject.

With regard to science learning in school, higher availability and usage of ICT and more unfair treatment of students by teachers, according to students' reports, were linked with students' decreased odds of belonging to the high-achieving group in science after accounting for all the other predictor variables. With every extra unit in both the *ICT availability at school* and *teacher fairness* indices, students were 10% less likely to be high achievers in science. While no causal links of ICT availability at school and teacher fairness with high achievement in science can be established here, meaning that decreased levels of ICT availability at school or increased teacher fairness would not necessarily lead to higher chances of high achievement, it is noteworthy that these two variables retained their statistical significance after other important variables, such as student sex and socioeconomic status, were taken into account.

The predictor variables that were included in the model explained a considerable proportion of the variance in high achievement in science at both levels of the analysis. Specifically, the final model (step 5) accounted for 50.6% of the between-student (level-1) and 83.6% of the between-school (level-2) variance in high achievement in science.

Discussion

The analysis of PISA 2012 and 2015 data reported in this paper provided some insights into potential predictors of high achievement in mathematics and science among 15-year-old students in Ireland. In light of earlier findings in the area, the most consistent findings of this study and their implications for policy and practice are discussed.

Overall, the findings of this study indicate that students' families might play a crucial role in enhancing students' chances of being high achievers in mathematics and science. Hence, it seems reasonable to suggest that further efforts to enhance collaboration between teachers, schools, and parents are needed, while there is also a need for educational policy to focus teaching and public awareness on the value of non-cognitive factors and wellbeing as well as of appropriate leisure activities for improving achievement. Along these lines, preschool practitioners, teachers, and school principals could raise parents' awareness of their important role as learning partners in shaping their children's learning on an ongoing basis from the early years. Broadly speaking, greater attention could be placed on initial preschool practitioner and teacher education and continuing professional development programmes to preparing these professionals as well as school principals to work in partnership with parents. Specific ways of how these could be materialised with reference to student- and family-level factors that were found to predict high achievement in mathematics and science in this study are discussed below. In the interest of clarity, the discussion of findings mirrors the approach adopted in the model-building process (see Fig. 1).

Student and family demographics

Student sex

In comparison with sex differences noted in overall achievement in mathematics and science in Ireland (e.g., Clerkin et al., 2015; Shiel et al., 2016), sex differences in high achievement in the two subjects at the bivariate level, as identified by this study, were larger. In general terms, this corroborates earlier findings in the area (e.g., Ellison & Swanson 2010; Gilleece et al., 2010; Stoet & Geary, 2013) indicating that the sex gap in mathematics and science among high-achieving students favouring males is consistently larger compared to the respective gaps in overall achievement in the two subjects. After taking other variables into account, though, sex differences in high achievement remained relatively small, following the unsystematic patterns found in research studies on overall achievement (e.g., Zhou et al., 2017).

These findings do not necessarily undermine the practical significance of sex differences in the context of high achievement; rather, they point towards potential factors on which the Irish education system could focus to address these differences. In the mathematics model, the gap between males and females in high achievement was no longer significant after student self-beliefs, dispositions, drive, and engagement were considered. This suggests that these variables may have accounted for sex differences among high achievers, indicating that males and females with equivalent self-beliefs, dispositions, drive, and engagement tended to be equally likely to be high achievers in mathematics. Based on these findings, it seems that one possible way to deal with sex differences in high achievement in mathematics in Ireland would be to focus on students' non-cognitive attributes, especially those of females.

It is also noteworthy that although sex differences were not large or systematic, they were in favour of males in both mathematics and science, suggesting that female students may be lagging behind in high achievement in the two subjects. Given that high achievement in these areas during schooling is associated with higher participation in STEM-related courses at third level, a way of bridging this sex gap would be for schools to assure that all students but, especially, females are aware of the range of STEM courses and careers that are available to them from an early age. However, as it was highlighted in the STEM Education Review Group report, "one key barrier in this regard [under-representation of females in the STEM workforce in Ireland] arises from the fact that, while parents are the main influencers when it comes to advising their daughters on how to define educational and career paths, they generally lack information about career options." (The STEM Education Review Group, 2016, p. 8). Hence, schools could cooperate with parents to make sure that the latter are in a position to make STEM careers and their associated benefits more desirable for female students, in particular, without, of course, neglecting male students in these efforts. Making STEM careers more desirable for all students but, especially, females falls under the key priorities of the current STEM Education Policy in Ireland (DES, 2017b). This study highlights the importance of such policies and of closely monitoring sex differences in high and overall achievement in all the subjects, as an important aspect of educational equity.

Socioeconomic status

One factor that yielded large differences between the two performance groups in both subjects was students' family socioeconomic status. Prior to and even after accounting for a range of variables, students coming from households with higher socioeconomic status were more likely to be high achievers in mathematics and science. This finding suggests that aspects of post-primary students' lives pertaining to their home possessions (e.g., books at home), and their parents' occupation and education are particularly important when it comes to their probabilities of belonging to the high-achieving groups in mathematics and science.

This finding corroborates those of the few studies that have examined high achievement in mathematics and science. For instance, studies by Gilleece et al. (2010) and Tourón et al. (2018) have shown that higher socioeconomic status is associated with higher chances of high achievement. Further, it echoes findings from a body of research that investigated the relationship between family socioeconomic status and overall achievement in a range of subjects, including mathematics and science (e.g., Reardon, 2011). This is an important finding in the context of educational equity, especially when the relationship is examined while other variables are taken into account, as it indicates that over and above anything else that is going on in students' lives as well as in their classes and schools, the socioeconomic status of their family is still an important predictor of their academic outcomes even at the highest levels of performance. This is despite significant efforts that have been made by the Department of Education in Ireland over the years to support the more socioeconomically disadvantaged (e.g., Delivering Equality of Opportunity in Schools [DEIS] initiative). Consequently, this finding begs questions for policymakers and teachers alike about what more can be done to deal with this important aspect of equity in education.

Although schools alone might not be able to decrease socioeconomic inequalities per se, their role in preventing such inequalities from shaping students' learning and outcomes is crucial. Along with existing initiatives that primarily target socioeconomically disadvantaged schools, efforts could also be made at the individual level. Specifically, to support learning among disadvantaged students, schools need to target their efforts to improve communication with parents in the most disadvantaged homes, and help develop home environments conducive to learning in mathematics and science. Parents from socioeconomically disadvantaged families should be made aware of the importance of the availability of adequate educational resources at home (e.g., books and appropriate space to study) for their children's learning as well as of useful resources that they could use at home. Additionally, they should be supported by teachers and schools in getting involved in their children's learning, and in encouraging their children to take part in extracurricular activities related to mathematics and science that are organised by the school or other bodies. Such practices could contribute towards mitigating the impact of inequalities on students' outcomes. Alongside such efforts, there must also be a stronger emphasis on identifying and nurturing talent of all students but especially of those coming from socioeconomically disadvantaged backgrounds. This could involve a stronger role for special education teachers in Irish schools, who currently seem to focus on addressing the needs of at-risk low-achieving students.

Student self-beliefs, dispositions, drive, and engagement

A range of variables related to students' self-beliefs, dispositions, drive, and engagement emerged as important predictors of high achievement in mathematics and science. This is not surprising given cognate research, such as Bandura's (1997) work on the importance of such non-cognitive constructs and, especially, self-beliefs in the context of the *social-cognitive theory* as well as Stankov's (2013) *predictability gradient hypothesis*. Such non-cognitive student attributes are likely malleable and responsive to change through appropriate schooling and interventions and, thus, merit consideration.

Overall, more positive self-beliefs and dispositions and greater drive and engagement were linked with higher chances of high achievement in mathematics and science. Student self-beliefs and, more specifically, domain-specific self-beliefs were among the strongest predictors of high achievement in mathematics and science in both the bivariate and the multilevel binary logistic regression analysis, with self-efficacy yielding the largest differences. The bivariate analysis indicated that self-efficacy yielded larger effect sizes for high achievement in mathematics compared to high achievement in science. Lee and Stankov (2018), who analysed, amongst others, PISA data to investigate the role of non-cognitive factors in overall mathematics achievement, rather than high achievement, also indicated that students' self-efficacy was the strongest predictor of achievement, even after accounting for home possessions and parental education as proxies for students' socioeconomic status. This study echoed and extended this finding to high achievement in both mathematics and science. This highlights the importance of retaining and, potentially, increasing the current emphasis on initiatives that purport to enhance these students' self-beliefs in mathematics and science, such as those established following the Junior Cycle wellbeing guidelines in Ireland (National Council for Curriculum and Assessment [NCCA], 2017).[7]

There were also variables related to students' self-beliefs, dispositions, drive, and engagement (e.g., enjoyment of science, expected occupational status) that yielded particularly large differences between the two performance groups in mathematics and science in a bivariate context but, when examined in a multiple regression context, were no longer significant predictors of high achievement. Given that most of the existing studies examined the relationships of these variables with high achievement in mathematics and science in a bivariate context (see, for example, Tourón et al., 2018), the findings of the

[7] Some caution is warranted in the interpretation of the relationship between students' self-efficacy (and how the construct is measured in PISA) and student achievement, including high achievement. In PISA, self-efficacy in mathematics and science could be described as a proxy for achievement in each of the two subjects, as students are asked to indicate their level of confidence in performing various types of mathematics and science tasks. Hence, its relationship with achievement could be interpreted as achievement explaining achievement, with those who think that they are able to perform complex tasks in the two subjects being the highest achievers.

current study extend existing research by taking into account other important student- and school-level variables in a multiple and multilevel regression context.

There were also a small number of instances where differences between the two performance groups in each subject in the bivariate analysis were in favour of one group but, when examined in a multiple regression context, were in favour of the other. Tourón et al. (2018) showed that, in a bivariate context, high-achieving students had higher levels of environmental awareness compared to their low-achieving peers, which was also the case in this study. However, when students' environmental awareness was included in the multilevel model of the PISA 2015 data in this study, it was found to be a negative predictor of high achievement in science. This suggests that students who were identical in all of the other examined variables but had a more thorough understanding of environmental issues were less likely to be high achievers in science. It is plausible that students with high knowledge of science have a greater awareness of the difficulties associated with reversing environmental problems. In any case, this is a finding that raises more questions than it answers, prompting further investigation of the underpinning mechanisms behind this relationship.

These findings provide a strong rationale for the current emphasis that the Irish education system has placed on students' self-beliefs and other non-cognitive attributes, as integral parts of their wellbeing. They also suggest that preschool practitioners, primary and post-primary teachers, and schools along with parents could work together to help students develop and strengthen such self-beliefs, which, in turn, are expected to work positively towards raising students' chances of becoming high achievers in mathematics and science.

Research and relevant initiatives in Ireland and elsewhere acknowledge wellbeing as a key issue in the holistic development of children and young people. The promotion of students' self-efficacy, as an important aspect of their wellbeing, is among the emphasised areas during early childhood education as well as at primary and post-primary levels in an ever-increasing number of countries around the world, including Ireland; for example, students' wellbeing is a central aspect within *Aistear*, the early childhood curriculum framework in Ireland (NCCA, 2009), the proposed revised primary curriculum (NCCA, 2020), and the curriculum frameworks for the Junior and Senior Cycles (i.e., post-primary education) (NCCA, 2011, 2017). This emphasis, which is justified by the findings of this study but also by data from PISA 2018[8], acknowledges the important role of teachers, schools, and parents in enhancing students' interest and engagement in learning, by ensuring that students develop positive perceptions about their abilities that, in turn, can help them improve and sustain their gains in achievement.

In Ireland, this emphasis has also led to the creation of a wide range of guidelines, innovative courses, and other resources (for examples, see the NCCA's website) to assist preschool practitioners, teachers, schools, and parents in their efforts to enhance students' wellbeing. Specifically, NCCA's guidelines include relevant information on how preschool practitioners, teachers, and school principals can work together with parents to promote children's wellbeing, including their self-beliefs. For instance, such a

[8] The PISA data for Ireland in 2018 indicated that students reported lower levels of life satisfaction compared to the OECD average, fewer students felt cheerful, joyful or proud than students, on average, across OECD countries, students often compared themselves to others, and over half of students reported that they tend to worry about what would happen if they fail an exam or test, indicating relatively poor wellbeing among students in Ireland compared to their counterparts in other countries (McKeown et al., 2019).

collaboration could be facilitated by individuals who serve as connection points between students' homes and their schools and who can gain insights into individual students' home circumstances (such as the Home School Community Liaison Officers within the DEIS initiative in Ireland) and, thus, act as coordinators of action. However, a requirement for parents to be able to support their children in developing and improving their self-beliefs in relation to mathematics and science is that parents themselves have strong self-beliefs in their mathematical and scientific abilities. This highlights that part of the overall efforts in improving students' self-beliefs may also need to be the corresponding improvement of their parents' self-beliefs. This is an area, though, that has not been examined by the current study, and one that could be examined by future research.

On the whole, the current emphasis on students' wellbeing especially within early childhood education in Ireland should be retained and increased, given that these constructs tend to be more malleable during these early years and, thus, positive developments during these early years are likely to have long-term benefits for students.

Student mathematics- and science-related activities and schedule, parent dispositions and support

Students who, according to their parents, were more frequently engaged in science-related activities at the age of 10 were significantly more likely to be high achievers in the subject, again, highlighting the importance of early knowledge about and engagement with materials related to school subjects for future academic success. It is noteworthy that this finding, based on a set of PISA questions about parents' support for science learning in the middle childhood years, highlights not only the importance of parental encouragement, curiosity, and involvement in their children's learning, but also how these can contribute to subsequent academic outcomes. The review of the existing literature on high achievement in mathematics and science indicated that previous studies in the area had not explored the role of early skills and engagement with relevant materials in high achievement in either of the two subjects, rendering this one of the first studies to address this topic.

This finding highlights the crucial role of early years and parents' support in promoting students' high achievement. On this basis, policymakers and schools should aim to raise parents' awareness about behaviours and practices that stimulate students' progress and those that do not. For instance, early interventions and programmes targeted at parents of young children should provide resources and supports to assist parents in enhancing their children's early knowledge and skills (e.g., literacy, numeracy etc.) and engaging their children in mathematics- and science-related activities during the early years.

In recognition of the usefulness of the resources for mathematics and science education that have already been developed by various Irish educational agencies, including those charged with supporting teachers to implement Project Maths[9], it is imperative that teachers and parents are made aware of this range of supplementary programmes available for teaching these two subjects and for supporting their children at home, respectively. This is important as solutions to educational issues may not always require

[9] Initiative developed as a response to a series of identified potential difficulties with mathematics education in Irish post-primary schools, including Ireland's results in international large-scale assessments, involving the development of revised syllabi in both Junior and Leaving Certificate mathematics (DES, 2010).

newly developed resources but rather awareness and appropriate use of existing ones. Bringing such resources together and organising them thematically, such as efforts being made by *scoilnet.ie*, *sfi.ie*, and *smartfutures.ie* as well as making them readily available for teachers, parents, and students are likely to assist towards this end.

Class and school characteristics

None of the school-level variables that yielded effect sizes large enough to progress to the multilevel binary logistic regression analysis retained their statistical significance in the multilevel models once other variables were accounted for. This does not imply that the role of classes and schools in the context of high achievement should be neglected. In fact, as discussed above, considerable proportions of variance in students' high achievement in mathematics and science were attributed to between-school differences. The fact that consistent findings about predictors of high achievement in the two subjects at the school level were not detected may be linked with the limited availability of information about these contexts. In light of the findings of this study, though, it appears that teachers and schools can act as mediators and facilitators in tackling differences between high- and non-high-achieving students based on certain characteristics stemming from the students themselves and their families that were found to contribute to the prediction of high achievement in mathematics and science.

At this point, it should be acknowledged that a singular focus on high achievers is highly likely to result in increased variance in performance and, hence, increased inequalities among students (e.g., Ferreira & Gignoux, 2014). Hence, this study advocates that a stronger emphasis on meeting the needs of high achievers, as discussed above, should occur with continuing attention to the needs of low achievers as well. Not only is such an approach expected to enhance Irish students' achievement across the performance continuum, but it is also expected to improve equity within the Irish education system.

Limitations

There are some limitations underlying this study that should be acknowledged and taken into account in the interpretation of the findings. Firstly, even though results based on data from international large-scale assessments that draw on nationally representative samples have the potential to be scalable and transferable, the non-experimental nature of these data did not allow for the establishment of causal relationships among the examined variables (L. Cohen et al., 2017). Research that attempted to explore the potential of these data in examining causal effects (see Rutkowski & Delandshere, 2016) has indicated that methodological limitations prompt more cautious interpretations than those of strict cause-and-effect. Hence, any inferences about the relationships between the examined variables should consider that the relationships may be reciprocal.

Secondly, most of the measures of the contextual information about students, their parents, classes, and schools administered were based on self-reports. It is acknowledged that this could lead to self-report response bias, as respondents may, intentionally or unintentionally, have provided distorted responses. This may, in turn, have reduced the validity of the inferences from these measures (R. J. Cohen & Swerdlik, 2009). Relatedly, PISA uses self-report measures to capture aspects of teaching and learning such as student-teacher interactions and teaching practices. It has been aptly argued, however, that the multifaceted nature of teaching and learning may not be fully captured through a series of responses to self-report questionnaires (see Kaplan & Kuger, 2016) and that other additional measures such as direct observations could complement such self-report data. Consequently, despite the wealth of information that PISA collects, it is acknowledged that, as a secondary analysis of data, this study was somewhat limited in terms of the measures and variables available.

Recommendations for future research

This study has contributed to the literature on high achievement in mathematics and science at post-primary level in Ireland by providing a detailed examination of PISA data. Notwithstanding this contribution, the limitations of this study and aspects of the research problem that were not explored lead naturally to recommendations for future research.

The series of bivariate and multilevel binary logistic regression analyses reported in this paper uncovered a range of contextual characteristics stemming from the students themselves, their homes, classes, and schools that are associated with high achievement in mathematics and science. Nevertheless, experimental and/or longitudinal studies would allow for the detection of potential causal links between such contextual characteristics and students' high achievement and of developments or changes in the characteristics of the target population at both the individual and the group levels. This was not possible in this study due to the non-experimental and cross-sectional nature of the data. While longitudinal studies that track students' development from early childhood up to post-primary levels and even beyond exist in Ireland (e.g., Growing Up in Ireland), longitudinal studies with a much stronger focus on mathematics and science and the needs of high-achieving students than that of the current efforts are required.

Despite the wealth of contextual information collected by PISA, there are many important variables that are not captured. Employing qualitative approaches to compare the distinct characteristics of education systems, curricula, and teaching practices from other countries to the reality in Irish schools could shed some more light on aspects that are not measured by PISA, such as the classroom and school contexts. Such investigations could, also, be complemented with national and international large-scale assessment data. Research projects, such as the TIMSS 1999 Video Study, a study of eighth-grade mathematics and science teaching in seven countries that involved recording and analysis of teaching practices in more than one thousand classrooms (Hiebert et al., 2003; Roth et al., 2006) and a study by Lyons et al., (2003) that was based on the TIMSS Video Study and collected videotape evidence from a small number of post-primary classrooms in Ireland, constitute illustrative examples of such research.

Future studies could seek to replicate the findings reported here and include information from the Junior and Leaving Certificate examinations, as the major high-stakes examinations in Ireland, by matching these data with information from PISA and other national or international large-scale assessments. In this way, patterns of high achievement in the subjects of interest and the relationships of contextual characteristics with high achievement in these subjects within a high-stake and a low-stake assessment context could be compared. It should be acknowledged, however, that under the General Data Protection Regulation (GDPR), such endeavours are likely to have their own challenges. At a minimum, for instance, they will need to ensure that students taking the national and international assessments give their permission for their state examination results to be used in this way.

Conclusion

This study focused on high achievement in mathematics and science among 15-year-old students in Ireland. This is an area that has attracted considerable policy attention in Ireland in the past decade primarily in light of national and international large-scale assessment results indicating that the Irish education system may be lagging behind with regards to high achievement in these subjects. High achievers, especially those in STEM-related areas, have a unique contribution to make to the future social and economic wellbeing of countries; hence, affording more students the opportunity to perform at the highest levels in these subjects should constitute an important part of all educational agendas. This research is timely given that it provides much-needed evidence during a period of ongoing efforts to raise interest and improve academic performance within the realm of STEM education and that a comprehensive review of the 2011 Literacy and Numeracy Strategy (to also include Digital Literacy) is among the stated priorities for primary and post-primary education in the current Programme for Government (Department of the Taoiseach, 2020), which also includes a focus on raising achievement standards at the upper end of the performance distribution. It is anticipated that this evidence will provide further impetus towards this end and inform efforts to address the challenges in relation to high achievement facing the Irish education system.

List of abbreviations

DEIS	Delivering Equality of Opportunity in Schools
DES	Department of Education and Skills
GDPR	General Data Protection Regulation
IDB	International Database Analyzer
IEA	International Association for the Evaluation of Educational Achievement
NAMER	National Assessments of Mathematics and English Reading
NCCA	National Council for Curriculum and Assessment
NFQ	National Framework of Qualifications
OECD	Organisation for Economic Co-operation and Development
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
STEM	Science, Technology, Engineering, and Mathematics
TIMSS	Trends in International Mathematics and Science Study

Acknowledgements

The author is grateful to Dr Anastasios Karakolidis for his valuable assistance during multiple stages of this research and for proofreading the manuscript.

Authors' contributions

Not applicable.

Funding

This research was supported by the Irish Research Council Government of Ireland Postgraduate Scholarship. The funding body was not responsible for the design of the study, the collection, analysis, and interpretation of data or the writing of the manuscript.

Availability of data and materials

The PISA datasets analysed in this study are publicly available on the PISA website (https://www.oecd.org/pisa/data/).

Declarations

Competing interests

Not applicable.

Received: 1 March 2022 / Accepted: 30 August 2022

Published online: 20 September 2022

References

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic & Physiological Optics: The Journal of the College of Optometrists*, *34*(5), 502–508. https://doi.org/10.1111/opo.12131

Asparouhov, T. (2009). *Multilevel data/complex sample: Weighting*. Mplus Discussion. http://www.statmodel.com/discussion/messages/12/3975.html

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.

Bronfenbrenner, U. (1994). Ecological models of human development. In T. Husén, & T. N. Postlethwaite (Eds.), *International Encyclopedia of Education* (2nd ed., vol. 3, pp. 1643–1647). Pergamon Press.

Clerkin, A., Perkins, R., & Cunningham, R. (2015). *TIMSS 2015 in Ireland: Mathematics and science in primary and post-primary schools*. Educational Research Centre.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education* (8th ed.). Routledge.

Cohen, R. J., & Swerdlik, M. E. (2009). *Psychological testing and assessment: An introduction to tests and measurement* (7th ed.). McGraw-Hill.

Department of Education and Skills. (2010). *Report of the Project Maths Implementation Support Group*. Department of Education and Skills. https://assets.gov.ie/24637/ec517c86c4cd4792b29adda2d6326d17.pdf

Department of Education and Skills. (2011). *Literacy and Numeracy for Learning and Life: The national strategy to improve literacy and numeracy among children and young people 2011–2020*. Department of Education and Skills. https://assets.gov.ie/24520/defd56aec10946798ab2d32a42dc0d86.pdf

Department of Education and Skills. (2016a). *Action Plan for Education 2016–2019*. Department of Education and Skills. https://assets.gov.ie/24370/ec3df78b298e4574ab2d7c98f02450b5.pdf

Department of Education and Skills. (2016b). *Ireland's National Skills Strategy 2025*. Department of Education and Skills. https://assets.gov.ie/24412/0f5f058feec641bbb92d34a0a8e3daff.pdf

Department of Education and Skills. (2017a). *National Strategy: Literacy and Numeracy for Learning and Life 2011–2020 - Interim review: 2011–2016. New targets: 2017–2020*. Department of Education and Skills. https://assets.gov.ie/24960/93c455d4440246cf8a701b9e0b0a2d65.pdf

Department of Education and Skills. (2017b). *STEM Education: Policy statement 2017–2026*. Department of Education and Skills. https://assets.gov.ie/43627/06a5face02ae4ecd921334833a4687ac.pdf

Department of Education and Skills. (2018). *Chief Inspector's report-January 2013-July 2016: Excellence in learning for all*. Inspectorate of the Department of Education and Skills. https://assets.gov.ie/25245/9c5fb2e84a714d1fb6d7ec7ed0a099f1.pdf

Department of the Taoiseach. (2020). *Programme for Government: Our shared future*. https://assets.gov.ie/130911/fe93e24e-dfe0-40ff-9934-def2b44b7b52.pdf

Ellison, G., & Swanson, A. (2010). The gender gap in secondary school mathematics at high achievement levels: Evidence from the American Mathematics Competitions. *Journal of Economic Perspectives*, *24*(2), 109–128. https://doi.org/10.1257/jep.24.2.109

Ferreira, F. H. G., & Gignoux, J. (2014). The measurement of educational inequality: Achievement and opportunity. *World Bank Economic Review*, *28*(2), 210–246. https://doi.org/10.1093/wber/lht004

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE.

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*(1), 2–18. https://doi.org/10.1037/a0024338

Gilleece, L., Cosgrove, J., & Sofroniou, N. (2010). Equity in mathematics and science outcomes: Characteristics associated with high and low achievement on PISA 2006 in Ireland. *International Journal of Science and Mathematics Education*, *8*, 475–496. https://doi.org/10.1007/s10763-010-9199-2

Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). John Wiley & Sons, Ltd.

Goldstein, H., Browne, W., & Rasbash, J. (2002). Partitioning variation in multilevel models. *Understanding Statistics*, *1*(4), 223–231. https://doi.org/10.1207/S15328031US0104_02

Gonzalez, E. J. (2014). Calculating standard errors of sample statistics when using international large-scale assessment data. In R. Strietholt, W. Bos, J. E. Gustafsson, & M. Rosén (Eds.), *Educational Policy Evaluation through International Comparative Assessments* (pp. 59–74). Waxmann.

Government of Ireland. (2018). *Action Plan for Education 2018*. Government of Ireland. https://assets.gov.ie/24349/c5eadbca-68be4e8e90da28e36f377452.pdf

Greaney, V., & Kellaghan, T. (2008). *Assessing national achievement levels in education*. World Bank. https://openknowledge.worldbank.org/handle/10986/6904

Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., Chui, A. M.-Y., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterbeek, W., Manaster, C., Gonzales, P., & Stigler, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study (NCES 2003-013)*. U.S. Department of Education, National Center for Education Statistics. https://nces.ed.gov/pubs2003/2003013.pdf

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2018). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge.

IEA. (2021). *Help manual for the IEA IDB Analyzer (Version 4.0)*. https://www.iea.nl

Johansson, S. (2016). International large-scale assessments: What uses, what consequences? *Educational Research*, *58*(2), 139–148. https://doi.org/10.1080/00131881.2016.1165559

Kaplan, D., & Kuger, S. (2016). The methodology of PISA: Past, present, and future. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing contexts of learning: An international perspective* (pp. 53–73). Springer.

Kartal, S. K., & Kutlu, Ö. (2017). Identifying the relationships between motivational features of high and low performing students and science literacy achievement in PISA 2015 Turkey. *Journal of Education and Training Studies*, *5*(12), 146–154. https://doi.org/10.11114/jets.v5i12.2816

Kourti, S. (2019). Student engagement in inquiry-based learning: Cognition, behavior and affect. *Eleventh Congress of the European Society for Research in Mathematics Education*. Utrecht University. Utrecht, Netherlands. https://hal.archives-ouvertes.fr/hal-02410139

Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences*, *65*, 50–64. https://doi.org/10.1016/j.lindif.2018.05.009

Lyons, M., Lynch, K., Close, S., Sheeran, E., & Boland, P. (2003). *Inside classrooms: A study of teaching and learning*. Institute of Public Administration.

Mang, J., Küchenhoff, H., Meinck, S., & Prenzel, M. (2021). Sampling weights in multilevel modelling: An investigation using PISA sampling structures. *Large-Scale Assessments in Education*, *9*(6), https://doi.org/10.1186/s40536-021-00099-0

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *The Journal of Mathematical Sociology*, *4*(1), 103–120. https://doi.org/10.1080/0022250X.1975.9989847

McKeown, C., Denner, S., McAteer, S., Shiel, G., & O'Keeffe, L. (2019). *Learning for the future: The performance of 15-year-olds in Ireland on reading literacy, science and mathematics in PISA 2018*. Educational Research Centre.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

National Council for Curriculum and Assessment. (2009). *Aistear: The early childhood curriculum framework - Principles and themes*. National Council for Curriculum and Assessment. https://curriculumonline.ie/getmedia/484bcc30-28cf-4b24-90c8-502a868bb53a/Aistear-Principles-and-Themes_EN.pdf

National Council for Curriculum and Assessment. (2011). *Social, personal and health education: Curriculum framework - Senior Cycle*. National Council for Curriculum and Assessment. https://ncca.ie/media/2688/sphe_framework.pdf

National Council for Curriculum and Assessment. (2017). *Junior Cycle wellbeing guidelines*. National Council for Curriculum and Assessment. https://ncca.ie/media/2487/wellbeingguidelines_forjunior_cycle.pdf

National Council for Curriculum and Assessment. (2020). *Draft primary curriculum framework: For consultation - Primary curriculum review and redevelopment*. National Council for Curriculum and Assessment. https://ncca.ie/media/4456/ncca-primary-curriculum-framework-2020.pdf

OECD. (2014). *PISA 2012 technical report*. PISA, OECD Publishing. https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf

OECD. (2017a). *PISA 2015 assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving (revised edition)*. PISA, OECD Publishing. https://doi.org/10.1787/9789264281820-en

OECD. (2017b). *PISA 2015 technical report*. PISA, OECD Publishing. https://www.oecd.org/pisa/sitedocument/PISA-2015-technical-report-final.pdf

Perkins, R., & Clerkin, A. (2020). *TIMSS 2019 Ireland's results in mathematics and science*. Educational Research Centre.

Perkins, R., Shiel, G., Merriman, B., Cosgrove, J., & Moran, G. (2013). *Learning for life: The achievements of 15-year-olds in Ireland on mathematics, reading literacy and science in PISA 2012*. Educational Research Centre.

Pitsia, V. (2021). *Investigating high achievement in mathematics and science in Ireland: An in-depth analysis of national and international assessment data* [Dublin City University]. https://doras.dcu.ie/25255/

Pitsia, V., Lysaght, Z., O'Leary, M., & Shiel, G. (2022). High achievement in mathematics and science among students in Ireland: An in-depth analysis of international large-scale assessment data since 2000. *Irish Educational Studies*. https://doi.org/10.1080/03323315.2022.2061563

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. Murnane, & G. Duncan (Eds.), *Whither opportunity? Rising inequality and the uncertain life chances of low-income children*. Russell Sage Foundation Press.

Roth, K. J., Druker, S. L., Garnier, H. E., Lemmens, M., Chen, C., Kawanaka, T., Rasmussen, D., Trubacova, S., Warvi, D., Okamoto, Y., Gonzales, P., Stigler, J., & Gallimore, R. (2006). *Teaching science in five countries: Results from the TIMSS 1999 video study - Statistical analysis report (NCES 2006-011)*. US Government Printing Office.

Rutkowski, D., & Delandshere, G. (2016). Causal inferences with large scale assessment data: Using a validity framework. *Large-Scale Assessments in Education*, *4*. https://doi.org/10.1186/s40536-016-0019-1

Shiel, G., & Kelleher, C. (2017). *An evaluation of the impact of Project Maths on the performance of students in Junior Cycle mathematics*. Educational Research Centre.

Shiel, G., Kelleher, C., McKeown, C., & Denner, S. (2016). *Future ready? The performance of 15-year-olds in Ireland on science, reading literacy and mathematics in PISA 2015*. Educational Research Centre.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. SAGE.

Sommet, N., & Morselli, D. (2017). Keep calm and learn multilevel logistic modeling: A simplified three-step procedure using Stata, R, Mplus, and SPSS. *International Review of Social Psychology*, *30*(1), 203–218. https://doi.org/10.5334/irsp.90

Stankov, L. (2013). Noncognitive predictors of intelligence and academic achievement: An important role of confidence. *Personality and Individual Differences*, *55*(7), 727–732. https://doi.org/10.1016/j.paid.2013.07.006

Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of PISA data. *PloS One*, *8*(3), e57988. https://doi.org/10.1371/journal.pone.0057988

The STEM Education Review Group. (2016). *STEM Education in the Irish school system - A report on science, technology, engineering and mathematics (STEM) education: Analysis and recommendations*. The STEM Education Review Group. https://assets.gov.ie/25068/d5c86a91ac3b43869f827438f58d88c0.pdf

Tourón, J., López-González, E., Lizasoain Hernández, L., García San Pedro, M. J., & Navarro Asencio, E. (2018). Spanish high and low achievers in science in PISA 2015: Impact analysis of some contextual variables. *Revista de Educación*, *380*, 156–184. https://doi.org/10.4438/1988-592X-RE-2017-380-376

Veas Iniesta, A., López-López, J. A., Gilar Corbi, R., Miñano Pérez, P., & Castejón Costa, J. L. (2017). Differences in cognitive, motivational and contextual variables between under-achieving, normally-achieving, and over-achieving students: A mixed-effects analysis. *Psicothema*, *29*(4), 533–538. https://doi.org/10.7334/psicothema2016.283

von Davier, M., Gonzalez, E., & Mislevy, R. J. (2009). What are plausible values and why are they useful? *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments*, *2*, 9–36.

Woltman, H., Feldstain, A., Mackay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 52–69. https://doi.org/10.20982/tqmp.08.1.p052

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, *31*, 114–128. https://doi.org/10.1016/j.stueduc.2005.05.005

Zhou, Y., Fan, X., Wei, X., & Tai, R. H. (2017). Gender gap among high achievers in math and implications for STEM pipeline. *Asia-Pacific Education Researcher*, *26*(5), 259–269. https://doi.org/10.1007/s40299-017-0346-1

Publisher's Note