

RESEARCH

Open Access



Examining educational inequalities: insights in the context of improved mathematics performance on national and international assessments at primary level in Ireland

Anastasios Karakolidis , Alice Duggan, Gerry Shiel and Joanne Kiniry

*Correspondence:
anastasios.karakolidis@erc.ie
Educational Research Centre,
Dublin, Ireland

Abstract

Background: Evaluations of effectiveness of educational reforms are often based on the level of improvement in student performance from one cycle of a particular assessment to the next. However, improvements in overall performance do not necessarily translate to improved equality. Indeed, improvements that favour certain subgroups of students can exacerbate educational performance gaps and thus, inequality. This research examines changes in equality of mathematics achievement and subgroup performance differences in Irish primary school students over time. Ireland constitutes an interesting case study due to the introduction of a new National Literacy and Numeracy Strategy in 2011, the initial implementation of which has been linked to significant improvements in student mathematics performance. This paper aims to investigate whether these improvements have been accompanied by improvements in equality.

Methods: Using data from the Irish National Assessments of Mathematics and English Reading (NAMER) and the Trends in International Mathematics and Science Study (TIMSS) from the period before and after the introduction of the Strategy, the study examines (i) deviations in student scores, (ii) variability in achievement at student and school levels and (iii) performance gaps based on demographic and socioeconomic factors over time. Bivariate analyses and multilevel regression models were used to identify student- and school-level variables related to mathematics performance.

Results: The results showed a decrease in variability in students' mathematics performance after the introduction of the Strategy; this decrease was statistically significant only for TIMSS. Additionally, there was a considerable decrease in variance in mathematics performance attributed to between-school differences over time. These findings constitute evidence of increased equality as performance differences between students and schools tended to shrink. Regarding performance gaps and variance in mathematics performance explained by background characteristics, this study provided mixed results. In NAMER, subsequent to the introduction of the Strategy, the performance gaps and variance in mathematics achievement explained by selected demographic and socioeconomic characteristics decreased considerably. However, this was not the case for TIMSS.

Conclusions: The evidence provided by this study suggests that Ireland has made reasonable progress in addressing inequality. However, there is room for improvement, as a significant proportion of the variance in student mathematics performance is still explained by demographic and socioeconomic characteristics.

Keywords: Large-scale assessments, Mathematics, Equality

Background

In many countries, national and international large-scale standardized assessments play an important role in assessing the quality of education systems and in developing and evaluating educational policies (Greaney & Kellaghan, 2008; Johansson, 2016). Large-scale assessments can be used to describe and monitor achievement trends over time, explore factors that may be related to achievement and compare achievement across subgroups of students. International assessments facilitate cross-country comparisons, since all participating countries follow similar procedures (Kellaghan et al., 2009), with league tables, in particular, drawing an inordinate amount of attention (Mullis, Martin, & Loveless, 2016; Mullis, Martin, Foy, et al., 2016). National assessments, on the other hand, allow for more flexibility in terms of design, implementation and reporting and can help draw attention to issues related to a particular country's education system.

Evaluations of the progress or effectiveness of educational reforms are often based on the level of improvement in student performance from one cycle of a particular assessment to the next. However, improvements in overall student performance do not necessarily translate to improved equality. Indeed, if improvements in performance favour certain subgroups of students, they can lead to or exacerbate educational performance gaps and thus, inequality. Educational inequality is an important issue from both practical and ethical perspectives. It can impede economic mobility and prosperity (The World Bank, 2005), and may also be considered ethically unacceptable, particularly if high levels of inequality are due to factors beyond the control of individual students (Checchi et al., 2010; Fleurbaey & Peragine, 2013).

Equitable education requires that differences in student outcomes related to background, economic or social circumstances are minimal (OECD, 2018). Therefore, attempts to address educational inequality require the ability to characterize uneven distributions of learning and performance, and identify the factors which determine them (UNESCO, 2018).

Ferreira and Gignoux's (2014) framework for conceptualising and measuring inequality in education reflects the above principles by defining two different types of inequality: (i) inequality of educational achievement, which is simply the degree of variability in student performance and (ii) inequality of opportunity, which can be described as the extent to which students' academic performance is shaped by background, demographic, socioeconomic and other pre-determined characteristics. Van de Werfhorst and Mijs's (2010) description of inequality in learning as dispersion in test scores complements Ferreira and Gignoux's (2014) definition of inequality of achievement.

Broer et al.'s (2019) analysis of two decades of Trends in International Mathematics and Science Study (TIMSS) data from several countries (not including Ireland) focused on inequalities in mathematics and science performance, and suggested that improvements in overall performance tend to be accompanied by a narrowing of achievement

gaps related to socioeconomic inequalities. However, this pattern did not hold for all countries; some countries that have seen significant improvements in overall mathematics performance (e.g., Singapore and Iran) and science (e.g., Lithuania) in TIMSS during the last 20 years have also faced a substantial widening of performance gaps between students of high and low socioeconomic status (Broer et al., 2019; Mullis, Martin, & Loveless, 2016). In these countries, improvements in student achievement were concentrated in certain groups of the population, namely students from socioeconomically advantaged families. In fact, based on Broer et al.'s (2019) results, only a few countries managed to improve their overall national performance in mathematics and science respectively, while also systematically reducing inequalities by (i) reducing the achievement gap between socioeconomically advantaged and disadvantaged students and (ii) improving the performance of students coming from families with low socioeconomic status.

Using data from 19 nationally representative studies, Reardon (2011) indicated that the achievement gap between students from high and low socioeconomic status families in the United States is much larger among children born in 2001 compared to children born 20 or more years earlier. This was mainly attributed to the strengthening association between family income and children's academic achievement, especially for families above the median income level. Reardon's (2011) study provided an illustrative example of how inequalities in society can have a corresponding impact on educational inequalities, as it examined patterns in academic achievement for a period in time when there was a significant rise in income inequalities.

In terms of inequality of achievement, namely the distribution of student scores, Freeman et al.'s (2010) analysis of TIMSS data from 1999 and 2007 revealed a positive relationship between a country's average performance and the within-country dispersion of scores. In other words, higher average performance at country-level was usually associated with lower levels of inequality in student achievement, within each country. The authors called this a "virtuous' equity-efficiency trade-off" (p. 2). Freeman et al.'s findings were corroborated by Mullis, Martin, and Loveless's (2016) analysis of TIMSS scores from 1995 and 2015. Additionally, focusing on the most recent TIMSS 2011 and 2015 cycles, Mullis et al. found that improvements in TIMSS international averages (across subjects and grade levels) were fairly evenly distributed at the highest and lowest levels of performance. The authors noted that "whether this is the beginning of a new trend remains to be seen..." (p. 67), suggesting that more research is required to confirm whether these patterns continue.

Achieving equality of academic outcomes, independently of students' nationality, gender, socioeconomic status and other factors beyond students' control is a key challenge faced by many education systems worldwide. Due to increasing rates of international migration (United Nations—Department of Economics and Social Affairs—Population Division, 2019), more education systems are seeing greater numbers of migrant students. Therefore, systems must be prepared for challenges related to issues of access and inclusion, particularly in countries that have seen very large increases in rates of immigration. Andon et al. (2014) used data from TIMSS, the Programme for International Student Assessment (PISA) and the Progress in International Reading Literacy Study (PIRLS) for 34 OECD countries to examine achievement gaps between migrant and native students.

They found consistent achievement gaps in favour of native students for reading, mathematics and science. The gap for science was slightly larger than those for mathematics and reading, which were empirically identical. Andon et al. note that background characteristics other than migrant status (e.g., socioeconomic status and native language)¹ may place some migrant students at an academic disadvantage.

Despite efforts to enhance equality in education, student background characteristics, especially socioeconomic status, continue to be systematically related to academic performance in most countries (Mullis, Martin, Foy, et al., 2016; OECD, 2019). When changes in overall student performance are observed, it is worthwhile to examine whether these changes are accompanied by greater equality—for example, by investigating patterns in achievement gaps between subgroups on variables such as socioeconomic status across different studies over time. This approach provides a more in-depth description of uneven distributions of learning and performance, as well as the factors which are associated with them – both of which are necessary in order to better understand the issue of educational inequality (UNESCO, 2018).

The current study

Although most countries participating in large-scale assessment programmes, such as TIMSS, aim to improve equality by improving performance among socioeconomically disadvantaged students, Broer et al. (2019, p. 3) argue that “quantifiable measures of these efforts are still lacking, especially those that are straightforward and easy to understand”.

This study aims to provide an enhanced approach to investigating how changes in performance are distributed across subgroups. Specifically, the study draws on national and international assessment data for Ireland to examine improvements in equality of achievement and alleviation of subgroup performance differences. Ireland constitutes an interesting case study, as in 2011, the Department of Education and Skills announced a new National Literacy and Numeracy Strategy 2011–2020 (Department of Education & Skills, 2011). The impetus for the Strategy was the large decline in reading literacy and mathematics scores among students in Ireland on PISA 2009 compared with earlier cycles (Perkins et al., 2012), though the extent of those declines has since been questioned (Cosgrove & Cartwright, 2014). The Strategy aimed to improve literacy and numeracy among children and young people at all levels in the education system, through such measures as increased instructional time, enhanced teacher preparation and a stronger focus on learning outcomes and analysis of achievement data in schools (Department of Education & Skills, 2011).

Another important aim of the Strategy was to alleviate educational disadvantage and reduce performance gaps attributed to student background, thereby creating greater equality in schools. There was a general emphasis on targeted support for students at greater risk of low performance (e.g., students from socioeconomically disadvantaged backgrounds) “because of the enormous impact improvement can have on the life-chances of these young people and also because it fosters greater equity in the

¹ Andon et al.'s (2014) findings apply to aggregated performance across OECD countries. Country-level analyses were not conducted.

education system and society in general” (Department of Education & Skills, 2011, p. 62). The Strategy put particular emphasis on disadvantaged schools, commonly referred to as DEIS schools. DEIS schools are those that are eligible to participate in the Delivering Equality of Opportunity in Schools (DEIS) plan in Ireland, which is the Department of Education’s main programme for the alleviation of educational disadvantage (Department of Education & Science, 2005; Department of Education & Skills, 2017a). Disadvantaged schools participating in DEIS are allocated resources including additional teaching staff, additional funding for resources, and access to supplementary programmes such as Reading Recovery. These schools’ resources and obligations related to DEIS are in addition to those associated with the Strategy.

The Strategy also aimed to bridge the performance gap between native and migrant students, including those whose first language is not the language of the school. This is an issue of growing importance in Ireland, as numbers of migrants have significantly increased over the last two decades (United Nations—Department of Economics and Social Affairs—Population Division, 2019).

After the introduction and initial implementation of this Strategy, significant improvements were observed in students’ performance on national and international assessments (Clerkin et al., 2016; Perkins et al., 2013; Shiel et al., 2014, 2016). However, the question remains to what extent these improvements were accompanied by decreases in inequality of achievement and performance gaps attributed to demographic and socioeconomic factors. Despite the existence of some international evidence, research on this topic specifically focused on the Irish context is limited. Some analysis has been conducted on the factors explaining mathematics and reading achievement on national and international studies in Ireland, without specifically investigating issues around equality (Cosgrove & Creaven, 2013; Kavanagh et al., 2015). Karakolidis et al. (2021) attempted to shed some light on this research problem by using National Assessment data from 2009 and 2014 to compare inequalities in 6th class mathematics and reading achievement, before and after the initial implementation of the Strategy. The results indicate that the improvements in overall pupil performance were accompanied by reduced inequalities. While all examined groups of pupils saw improvements in both reading and mathematics over time, the improvements particularly favoured groups of pupils who had lower performance than their counterparts in 2009, leading to smaller performance gaps in 2014. These findings were consistent for reading and mathematics. Despite the evidence of reduced inequalities provided by this study, there has been no in-depth investigation of the impact of the Strategy on equality among different groups of students, analysing data from multiple studies, grade levels and time points.

The current study examines whether improvements in Irish primary school students’ overall mathematics performance, after the introduction and initial implementation of the Strategy, are accompanied by improvements in equality. Data from the National Assessments of Mathematics and English Reading (NAMER) and TIMSS are used to cover the period before and after the introduction of the Strategy. Drawing on Ferreira and Gignoux’s framework on inequality of achievement and opportunity, the study examines (i) deviations in student scores, (ii) variability in achievement at student and school levels and (iii) performance gaps based on demographic and

Table 1 Sample sizes in NAMER 2009, 2014 and TIMSS 2011, 2015 in Ireland

	NAMER		TIMSS	
	2009	2014	2011	2015
Number of schools	140	139	151	149
Number of students	3832	3312	4560	4344

In NAMER 2014, a total sample of 4144 6th-grade students completed the mathematics assessment. However, experimental test booklets comprising fewer items were administered to a subsample. To facilitate a direct comparison of mathematics scores across the years, data for students who completed the shorter version of the mathematics test were not included in the analyses reported in this paper

socioeconomic factors over time. Additionally, the study explores whether and to what extent evidence from different studies and grade-level cohorts yields consistent results, where equality is concerned.

Methods

Data

This paper presents results from secondary analyses of national and international assessment data for primary school-aged students. Data from NAMER 2009, NAMER 2014, TIMSS 2011 and TIMSS 2015 were used. NAMER is a cross-sectional study that takes place every 5/6 years in Ireland, assessing 2nd- and 6th-grade students' performance in mathematics and reading. TIMSS is a study of the International Association for the Evaluation of Educational Achievement (IEA) that takes place every four years and assesses 4th- and 8th-grade students' mathematics and science achievement. Both NAMER and TIMSS involve the administration of curriculum-based assessments, as well as questionnaires to gather contextual information. The current study focuses on the primary education context using mathematics achievement data for students in 4th grade (10 years old, on average in TIMSS) and 6th grade (12 years old, on average in NAMER).

Participants

NAMER and TIMSS have similar complex designs, selecting a representative sample of schools and students within a country. In each cycle of NAMER and TIMSS, a two-stage cluster sampling methodology is utilized; a representative sample of schools is selected using stratified sampling and based on probability proportional to size, and then a number of classes within each school are randomly selected to participate, with all students in selected classes asked to take part. In both NAMER and TIMSS, a maximum of two classes per grade level are selected to participate. However, it is common for schools to have only one class per grade level; therefore, in the majority of the cases only one class per school participates in the studies. Table 1 presents the sample sizes for each cycle of NAMER and TIMSS used in the current study.

Table 2 (below) provides more information regarding the background characteristics of students participating in each cycle of NAMER and TIMSS used in the current study.

Measures

Mathematics achievement constitutes the main outcome measure in this study. The NAMER mathematics assessment is based on the current Primary School

Mathematics Curriculum in Ireland (Department of Education & Science, 1999). The assessment comprises items classified by content areas (number, algebra, shape & space, measures and data) that require specific process skills (recall, understand, implement, integrate, connect, reason, apply and problem-solve) (Educational Research Centre, 2009). Item parameters and student scale scores were estimated iteratively within the IRT scaling software Bilog-MG (Zimowski et al., 1996) with the final parameters and scores representing the “best fit” solution. In 2009, the overall mean score and standard deviation were set at 250 and 50, respectively.

The assessment framework for the TIMSS mathematics test is also organized in terms of content domains and cognitive domains (processes). At 4th grade, the content domains are number, geometric shapes & measures and data display. The cognitive domains are knowing, applying and reasoning. Like in NAMER, each item in the test is classified as belonging to one content domain and one cognitive domain (Mullis & Martin, 2013). Overall student performance in TIMSS is reported with reference to an IRT-based scale with an international country centerpoint (mean) of 500 and a standard deviation of 50 that was established in the first TIMSS cycle in 1995. Subsequent iterations of the NAMER and TIMSS tests were linked to the initial scales allowing for comparability of performance over time.

While there are broad similarities in terms of content and process between NAMER and TIMSS, countries in TIMSS, including Ireland, have relatively little control over the content of the items selected, though an international test-curriculum matching analysis revealed that the content and processes underpinning the vast majority of mathematics items had been covered by 4th-grade students in Ireland (Mullis & Martin, 2013). Both NAMER and TIMSS employ a rotated block design, with individual students completing a proportion of the full item pool. The data used in this paper were gathered using paper-based assessments.

In addition to measuring students’ achievement, both NAMER and TIMSS involve the administration of questionnaires to students, parents/guardians, teachers and school principals, to collect contextual and background information. These questionnaires are administered in conjunction with the tests and can be linked to achievement data.

To investigate equality in student performance across years and subjects, several demographic and socioeconomic factors were examined using data from the questionnaires (see Table 2). Variables related to educational possessions (i.e., number of books at home) and parents’ level of education were used as proxies for the socioeconomic status of the family. Country of birth and measures of the language most often spoken at home were both used to provide information relevant to students’ migration status.² It should be noted that data on parents’ education were not collected in NAMER 2009, nor were data on students’ country of birth collected in TIMSS 2011.

Schools’ disadvantaged (i.e., DEIS) status was used as a socioeconomic indicator at school level in both cycles of NAMER and TIMSS (Eivers & Clerkin, 2012; Shiel et al., 2014). At the time that the Strategy was published, schools were designated as

² It has been noted across several cycles of national and international assessments that, of the students that reported having been born outside of Ireland, a considerable proportion were born in Northern Ireland or other parts of the United Kingdom, and were therefore native English speakers. Thus, students’ country of birth was considered alongside the language most often spoken at home, to provide more accurate information on migration status.

Table 2 Demographic features of the samples (%)

Variables	Categories	NAMER 2009	NAMER 2014	TIMSS 2011	TIMSS 2015
Gender (biological sex)	(0) Girl	48.3	53.6	49.5	47.3
	(1) Boy	51.7	46.4	50.5	52.7
Country of birth	(0) Outside of Ireland	15.0	11.7	–	8.8
	(1) Ireland	85.0	88.3	–	91.2
Language most often spoken at home ^a	(0) Another language	5.4	6.9	15.8	12.5
	(1) English/Irish	94.6	93.1	84.2	87.5
Number of books at home	(0) 100 or fewer	56.5	54.5	60.4	62.6
	(1) More than 100	43.5	45.5	39.6	37.4
Highest level of parent education	(0) No third-level education	–	38.5	41.2	30.8
	(1) Third-level education	–	61.5	58.8	69.2
School disadvantaged status	(0) Non-disadvantaged	78.2	79.2	80.9	81.9
	(1) Disadvantaged	21.8	20.8	19.1	18.1

Weighted percentages are presented. Missing cases were excluded

^a The questions asked about language spoken at home differed in each study. NAMER asked students to report whether or not they speak the language of the test (English/Irish) or a different language at home. TIMSS used a Likert scale asking students to report how often they speak the language of the test at home (Always, Almost Always, Sometimes, Never). The percentages presented above in the “Another language” category for TIMSS show the proportion of students reporting that they sometimes or never speak the language of the test at home

disadvantaged with reference to an index created in 2006 based on the proportions of students in a school, as reported by the school principal, for each of six equally-weighted factors: unemployment, lone parenthood, membership of the Traveller community, family size, access to free book grants, and use of local authority accommodation (Archer & Sofroniou, 2008; Kavanagh et al., 2017).³ When allocation to DEIS was announced in 2006, 197 schools were allocated to Band 1 (the most disadvantaged schools, which received greater levels of support), and 141 to Band 2 (less-disadvantaged schools) (Department of Education & Skills, 2017b). For the purposes of the current study, schools in both DEIS categories were grouped together.

Table 2 presents the variables examined in this study along with their respective categories and the percentages of students belonging to each category. All variables that were not already dichotomous were recoded into binary variables to facilitate comparisons of performance gaps.

Analysis

To examine inequalities in student mathematics performance associated with demographic and socioeconomic factors over time and across studies, statistical analysis was conducted in two stages, for both NAMER and TIMSS. In the first stage, bivariate analyses were performed to identify which of the examined student- and school-level variables were statistically significantly related to student performance, as well as to indicate the magnitude of the performance gaps; Hedge’s *g* effect sizes were computed.

³ Following the publication in 2017 of a new Action Plan for DEIS (Department of Education and Skills, 2017a), a new method of identifying disadvantaged schools was introduced. Drawing on small-area data collected in five-year national censuses of the population, the new approach uses the average level of socioeconomic disadvantage in the immediate area of each student’s home to define school-level disadvantage. All DEIS schools whose data contributed to the current study were designated prior to 2017.

In stage two, the extent to which each factor contributed to the explanation of students' mathematics performance, after accounting for the other variables, was examined. The aim was to measure the variance in achievement explained by demographic and socioeconomic factors and to compare explained variance across studies and cycles. The examined factors were included as explanatory variables in a series of multilevel linear regression models of student achievement in mathematics.

Multilevel analysis was applied to allow the investigation of the variance in achievement at both student and school levels, while accounting for the clustered nature of the samples in both studies (i.e., NAMER and TIMSS). With cluster sampling, students within the selected clusters (classes and schools) may be more similar to each other than they are to students in the target population in general. This violates an important assumption of many statistical models that cases in the sample are independent of each other. Lack of independence can lead to underestimation of standard errors, and subsequently increase the risk of a Type I error (Field, 2017). Multilevel models provide more accurate estimates by estimating the variation in the dependent variable that is attributable to differences within and between clusters (Tarling, 2009; Woltman et al., 2012). In this study, two-level analysis was applied with students at level one and schools at level two. A decision was made not to run three-level analysis, with classes in level two, as in most cases, only one class per school participated in the studies, meaning that, in such cases, the class level is identical to the school level. In addition, in Ireland, where there are two or more classes at a grade level, students are usually assigned at random to classes, something that further minimizes the between-class differences.

Models were run in multiple steps. In step 1, models included only background variables that were measured in both NAMER and TIMSS to allow for comparisons across studies. In step 2, models were constructed that were comparable across the two cycles of the same study, including variables that were measured in both cycles of each study. For NAMER 2014 and TIMSS 2015, step 3 models were constructed to further explore the role of variables that had not been measured in the previous cycles of each study in explaining student mathematics achievement, after accounting for other demographic and socioeconomic factors. Finally, to draw clearer conclusions about the changes in the performance gaps across the years, data from both cycles were analysed in a single multilevel model for each of the two studies and the interactions between the year of the study and each one of the explanatory variables were tested for statistical significance.

Level-one slopes were allowed to vary across schools; however, they did not significantly improve any of the models. Also, intra- and cross-level interactions were examined for the final models of each cycle; statistically significant interactions are reported in the results section. Sampling weights were used at both levels. Specifically, following Rutkowski et al.'s (2010) guidelines, the combined student and class weights were used at level one, while the final school weights were used at level two. Replicate weights were also taken into account in the bivariate analyses.

The IBM SPSS Statistics 25 (IBM, 2017), the IDB Analyser 4 (IEA, 2019) and the Mplus 8 (Muthén & Muthén, 2017) software were used to perform the analyses.

Table 3 Average performance in mathematics on NAMER 2009 and 2014, and TIMSS 2011 and 2015

	Mean (SE)	SD (SE)
NAMER		
2009	250.0 (2.35)	50.00 (0.83)
2014	261.7 (2.07)	48.92 (0.87)
TIMSS		
2011	527.4 (2.60)	77.95 (1.62)
2015	547.3 (2.14)	73.13 (1.24)

Statistically significant differences between cycles in bold

Results

From 2011 onwards, following the introduction and initial implementation of the Literacy and Numeracy Strategy, Irish primary school students’ mathematics performance on NAMER and TIMSS improved significantly. Table 3 shows improvements in average mathematics performance on NAMER between 2009 and 2014 ($g=0.24$), and improvements on TIMSS between 2011 and 2015 ($g=0.26$). The magnitude of the changes across cycles for both NAMER and TIMSS, as indicated by Hedge’s g , would be classified as small to medium (Cohen, 1988); however, according to the What Works Clearinghouse framework (an initiative of the United States Department of Education), effect sizes greater than 0.25 can be viewed as substantively important in educational research (What Works Clearinghouse, 2020).

Alongside the increase in average student achievement, a small decrease in standard deviations was noted in both studies (Table 3). However, the reduced variability in scores, as indicated by the decrease in standard deviations, was statistically significant only for TIMSS.

In the sections that follow, the results of NAMER and TIMSS for key groupings of students are presented.

Descriptive and bivariate statistics

Analysis of NAMER data

Table 4 shows the mathematics performance gaps for NAMER 2009 and 2014 for selected groups of students.

In both cycles of NAMER, mathematics achievement was statistically significantly associated with the number of books in students’ homes and their schools’ disadvantaged status. In NAMER 2014, parental education, which was not measured in 2009, was another factor that was statistically significantly linked to mathematics achievement. Students’ gender, their country of birth and the language spoken at home were not statistically significantly related to mathematics performance in either cycle. For each of the statistically significant performance gaps, substantively important effect sizes were observed.

Students who had more than 100 books at home performed significantly better than their peers with 100 books or fewer in both 2009 ($g=0.70$) and 2014 ($g=0.56$). In 2014, students whose parents had completed some third-level education achieved a significantly higher mean mathematics score than students whose parents had completed no third-level education ($g=0.54$). Students in non-disadvantaged schools achieved a

Table 4 Differences in mathematics performance between NAMER 2009 and 2014 for selected groups of students

	Mean (SE)	SD	g
Gender			
2009			
Girls	247.0 (3.27)	48.37	0.12
Boys	252.8 (2.48)	51.32	
2014			
Girls	260.0 (2.46)	46.95	0.08
Boys	263.7 (2.55)	51.06	
Country of birth			
2009			
Ireland	251.1 (2.66)	49.61	0.07
Outside of Ireland	247.9 (3.82)	50.42	
2014			
Ireland	262.4 (2.14)	49.10	0.06
Outside of Ireland	259.4 (3.10)	45.95	
Language spoken at home			
2009			
English/Irish	251.4 (2.54)	49.57	0.22
Other	240.5 (5.10)	49.69	
2014			
English/Irish	262.3 (2.16)	48.87	0.07
Other	259.0 (3.42)	47.22	
No. of books at home			
2009*			
100 or fewer	237.6 (2.32)	46.93	0.70
More than 100	270.4 (3.35)	46.75	
2014*			
100 or fewer	252.0 (2.05)	46.88	0.56
More than 100	277.9 (2.25)	45.81	
Highest level of parent education ^a			
2009			
No third-level education	–	–	–
Third-level education	–	–	
2014*			
No third-level education	248.4 (2.28)	46.10	0.54
Third-level education	273.4 (2.04)	46.73	
School disadvantaged status			
2009*			
Disadvantaged	226.9 (5.53)	50.49	0.58
Non-Disadvantaged	255.1 (2.68)	47.93	
2014*			
Disadvantaged	246.8 (4.79)	50.32	0.39
Non-Disadvantaged	265.7 (2.21)	47.78	

* $p < .05$ ^a Data on parental education were not collected in NAMER 2009

significantly higher mean mathematics score than students in disadvantaged schools, in both cycles, with effect sizes of 0.58 in 2009 and 0.39 in 2014.

For each of these variables, improvements in overall mathematics performance particularly favoured groups of students who performed lower than their peers in NAMER 2009, leading to smaller achievement gaps in NAMER 2014. It is noteworthy that students in disadvantaged schools improved in their mathematics performance by almost 20 score-points on average between 2009 and 2014, whereas the improvement for students in non-disadvantaged schools was 10.6 score-points.

Analysis of TIMSS data

Table 5 shows the mathematics performance gaps for TIMSS 2011 and 2015 for selected groups of students.

In both cycles of TIMSS, mathematics achievement was statistically significantly associated with the number of books in students' homes, their parents' level of education and their school's disadvantaged status. The language spoken by students at home was statistically significantly associated with mathematics achievement in 2011, but not in 2015. Student gender and country of birth were not statistically significantly associated with mathematics achievement in either cycle.

In both 2011 and 2015, students who had more than 100 books at home performed significantly better than their peers who had 100 books or fewer, with effect sizes of 0.51 and 0.58 for each respective cycle. Similarly, in both cycles, students whose parents had completed some third-level education outperformed those whose parents had not completed any third-level education, with effect sizes of 0.56 in 2011 and 0.65 in 2015. The mean mathematics score of students in non-disadvantaged schools was significantly higher than that of students in disadvantaged schools in both 2011 ($g=0.58$) and 2015 ($g=0.52$). In 2011, students who spoke the language of the test (i.e., English/Irish) at home performed significantly better than students who spoke another language ($g=0.21$); a similar difference was observed in 2015.

As shown in Table 5, improvements in overall mathematics achievement in TIMSS 2015 were not evenly distributed across the examined groups of students. The performance of some groups of students with lower scores in TIMSS 2011 (e.g., students with fewer than 100 books at home and those whose parents had not completed any third-level education) did not improve to the same extent as their higher-scoring counterparts. These discrepancies in the size of improvements in performance led to some greater achievement gaps in TIMSS 2015. Only the effect size for schools' disadvantaged status decreased between 2011 and 2015.

Multilevel regression analysis

Multilevel regression results for NAMER

To further examine how the variables examined so far contribute to the explanation of students' performance in the NAMER mathematics tests, they were included as explanatory variables in a series of multilevel models. Firstly, the null models (without any explanatory variables) were run for both NAMER cycles. As indicated by the intra-class correlation coefficients (ICC), there was a substantial decrease in the proportion of

Table 5 Differences in mathematics performance between TIMSS 2011 and 2015 for selected groups of students

	Mean (SE)	SD	g
Gender			
2011			
Girls	525.7 (3.56)	74.93	0.04
Boys	529.1 (3.29)	80.75	
2015			
Girls	545.1 (2.60)	70.86	0.06
Boys	549.4 (2.85)	75.05	
Country of birth ^a			
2011			
Ireland	–	–	–
Outside of Ireland	–	–	
2015			
Ireland	547.5 (2.17)	72.67	0.01
Outside of Ireland	547.7 (6.75)	76.86	
Language spoken at home			
2011*			
English/Irish	530.9 (2.82)	76.88	0.21
Other	514.7 (5.00)	80.88	
2015*			
English/Irish	550.1 (2.34)	72.61	0.21
Other	534.8 (5.51)	76.93	
No. of books at home			
2011*			
100 or fewer	517.4 (2.68)	75.37	0.51
More than 100	555.0 (3.18)	69.60	
2015*			
100 or fewer	535.2 (2.45)	71.85	0.58
More than 100	575.7 (2.52)	65.20	
Highest level of parent education			
2011*			
No third-level education	508.8 (3.02)	74.45	0.56
Third-level education	549.7 (2.98)	71.06	
2015*			
No third-level education	519.8 (2.76)	73.41	0.65
Third-level education	564.8 (2.45)	66.91	
School disadvantaged status			
2011*			
Non-disadvantaged	491.5 (5.43)	74.52	0.58
Disadvantaged	535.9 (2.94)	81.80	
2015*			
Non-disadvantaged	516.7 (4.71)	70.01	0.52
Disadvantaged	554.1 (2.25)	78.93	

* $p < .05$ ^a Data on country of birth were not collected in TIMSS 2011

Table 6 Multilevel modelling of NAMER 2009 and 2014 mathematics achievement – Grade 6

Variables	Step 1 (comparison model across studies)		Step 2 (comparison model across cycles)			Step 3		Combined model	
	B (SE)	2009	2014	2009	2014	2009	2009–2014	B (SE)	2009–2014
<i>Student-level</i>									
Gender (male)	12.2 (2.62)*	4.9 (2.27)*	4.9 (2.22)*	–	5.1 (2.20)*	–	12.5 (2.7)*	–	12.5 (2.7)*
Books at home (more than 100)	23.6 (2.8)*	21.3 (2.15)*	21.1 (2.19)*	–	15.9 (2.22)*	–	22.2 (1.8)*	–	22.2 (1.8)*
Language spoken at home (English/Irish)	– 2.8 (4.13)	– 7.33 (3.52)*	– 8.2 (4.51)	–	– 9.6 (4.65)*	–	– 5.9 (3.2)	–	– 5.9 (3.2)
Country of birth (Ireland)	–	–	1.0 (3.91)	–	2.6 (3.97)	–	0.4 (2.6)	–	0.4 (2.6)
Parent education (third-level)	–	–	–	–	16.5 (2.16)*	–	–	–	–
Year (2014)	–	–	–	–	–	–	16.8 (5.4)*	–	16.8 (5.4)*
<i>School-level</i>									
Disadvantaged status (Disadvantaged)	– 20.0 (7.36)*	– 5.2 (6.48)	– 5.3 (6.49)	–	– 2.5 (6.47)	–	– 12.2 (5.1)*	–	– 12.2 (5.1)*
<i>Interactions</i>									
Year*Gender	–	–	–	–	–	–	– 7.7 (3.5)*	–	– 7.7 (3.5)*
Intercept	243.2 (6.3)*	261.5 (4.27)*	261.6 (4.47)*	–	253.3 (4.41)*	–	244.6 (5.3)*	–	244.6 (5.3)*
<i>Variance explained (R²)</i>									
Student-level	9.0%	5.8%	5.7%	–	8.7%	–	9.2%	–	9.2%
School-level	13.7%	2.0%	2.0%	–	0.5%	–	7.2%	–	7.2%
Overall	10.0%	5.4%	5.3%	–	7.8%	–	8.9%	–	8.9%

* *p* < .05

variance in students' mathematics performance that was attributed to between-school differences in NAMER 2014 ($\rho = 14.7\%$) compared to NAMER 2009 ($\rho = 23.2\%$). The multilevel model tables present the unstandardized coefficients (B), which are based on scaled scores, along with their respective standard errors (SE) for each one of the examined variables; the values refer to the categories in brackets. Table 6 presents the models for NAMER 2009 and 2014.

In the step-1 model for 2009 in Table 6, which shows only those background variables measured in both NAMER and TIMSS, gender, number of books in the home and school disadvantaged status were statistically significant. Specifically, boys, students with more than 100 books at home and those attending schools that are not disadvantaged were more likely to perform better on average than other students. Overall, these factors explained 10% of the differences in student mathematics achievement in NAMER 2009.

After including an additional variable (i.e., country of birth) in the second multilevel model for NAMER 2009 (step 2 in Table 6), gender, books at home and schools' disadvantaged status were again statistically significantly associated with students' mathematics achievement. Students' country of birth and the language spoken at home were not statistically significantly related to mathematics achievement. Schools' disadvantaged status was the strongest variable in this model, with students attending disadvantaged schools (i.e., schools participating in the DEIS plan in Ireland) performing, on average, 20.1 score-points lower in mathematics than their peers attending non-disadvantaged schools. The examined factors included in the step-2 model for mathematics 2009, again, explained 10% of the differences in student mathematics achievement.

As far as the multilevel regression results for NAMER 2014 are concerned, in the step-1 and step-2 models, which each explained just over 5% of the variation in mathematics performance, gender and number of books in the home were significant. Language spoken at home was significant in step 1, but not in step 2. The third model for 2014 mathematics (step 3 in Table 6) shows that, in addition to gender and books at home, parents' level of education (which was not measured in NAMER 2009) and language spoken at home were both statistically significant factors. Specifically, with other variables held constant, students with at least one parent who completed some third-level education were expected to perform better than other students by 16.5 score-points. Although in the bivariate analysis no significant differences between students speaking English/Irish and another language at home were detected, after taking into account other demographic and socioeconomic characteristics, students who did not speak the language of the test at home were expected to perform better in mathematics by 9.6 score-points. This effect is in the opposite direction to the non-significant differences in favour of speakers of the language of the test in the bivariate analyses. The third model explained 7.8% of the total variance in student mathematics achievement in NAMER 2014.

An examination of the step-2 models for 2009 and 2014 (which include the same explanatory variables) facilitates a direct comparison of inequalities attributed to socioeconomic and demographic factors across the two cycles. In both cycles, gender and

Table 7 Multilevel modelling of TIMSS 2011 and 2015 mathematics achievement—Grade 4

Variables	Step 1 (comparison model across studies)		Step 2 (comparison model across cycles)		Step 3		Combined model
	<i>B (SE)</i>		<i>B (SE)</i>		<i>B (SE)</i>		
<i>Student-level</i>	2011	2015	2011	2015	2011	2015	2011—2015
Gender (male)	9.2 (3.90)*	5.4 (3.25)	9.4 (3.62)*	5.6 (3.28)	–	6.0 (3.29)	7.6 (2.5)*
Books at home (more than 100)	29.9 (3.16)*	34.6 (3.16)*	22.0 (3.14)*	26.4 (3.49)*	–	26.3 (3.53)*	24.0 (2.4)*
Language spoken at home (English/Irish)	4.4 (4.83)	11.1 (5.74)	5.9 (4.99)	9.0 (5.5)	–	– 9.0 (8.75)	7.3 (3.3)*
Country of birth (Ireland)	–	–	–	–	–	– 31.0 (10.00)*	–
Parent education (third-level)	–	–	26.8 (3.77)*	32.2 (3.83)*	–	32.8 (9.96)*	29.4 (3.1)*
Year (2015)	–	–	–	–	–	–	14.8 (4.1)
<i>School-level</i>							
Disadvantaged status (Disadvantaged)	– 28.8 (8.75)*	– 27.4 (6.98)*	– 21.8 (8.64)*	– 23.0 (6.56)*	–	– 23.8 (6.50)*	– 22.5 (5.5)*
<i>Interactions</i>							
Language spoken at home*Country of birth	–	–	–	–	–	29.3 (9.96)*	
<i>Intercept</i>	520.5 (5.36)*	531.7 (6.13)*	505.6 (5.59)*	514.3 (6.30)*	–	533.5 (8.63)*	502.9 (4.5)*
<i>Variance explained (R²)</i>							
Student-level	5.1%	6.6%	8.5%	11.0%	–	11.5%	11.0%
School-level	15.7%	29.6%	9.9%	25.8%	–	26.7%	14.7%
Overall	6.6%	8.2%	8.7%	11.9%	–	12.4%	11.3%

* $p < .05$

number of books at home were significantly related to mathematics achievement. Language spoken at home was not significant in the step-2 models for either year, though it was significant in the step-3 model for 2014. Country of birth was not a statistically significant explanatory variable of mathematics achievement in either 2009 or 2014, after accounting for other factors in the model. School disadvantaged status, which was a significant factor in the bivariate analysis and the strongest explanatory variable in 2009, was not statistically significant in the 2014 model, indicating that its effects were explained by the student variables, gender and books in the home. Overall, the variance attributed to the examined performance gaps in 2014 ($R^2 = 5.3\%$) was about half the magnitude for 2009 ($R^2 = 10.0\%$). This finding indicates that the inequalities attributed to the examined factors were considerably smaller in 2014.⁴

In order to more precisely identify the factors that significantly contributed to the reduction in variance explained by these variables, data from both NAMER cycles were included in a single model and the interactions between the year (2009/2014) and each of the examined explanatory variables were tested, after controlling for the main effects. The results are summarized in Table 6 (Combined model); non-significant interactions

⁴ Intra- and cross-level interactions were examined for the final NAMER models for both 2009 and 2014, but none of them reached statistical significance.

were excluded from the model. As Table 6 shows, only the interaction between the year of the study and gender was statistically significant. The negative sign of the interaction confirms that, after controlling for other factors, the gender gap was significantly reduced in NAMER 2014. It should be noted that the rest of the interaction terms, even though non-significant, indicated in all cases reduced performance gaps in NAMER 2014 compared to 2009.

Multilevel regression results for TIMSS

Table 7 presents the multilevel regression results for TIMSS 2011 and 2015. A comparison of the level-two variance in mathematics achievement in TIMSS 2011 ($\rho = 17.7\%$) and 2015 ($\rho = 10.2\%$) indicated that, over time, there was a substantial decrease in the proportion of variance in 4th-grade students' mathematics performance attributed to between-school differences.

Similar to NAMER results for 6th-grade students, the multilevel models for TIMSS 2011 results indicated that gender, books at home, parents' education and schools' disadvantaged status were significantly related to 4th-grade student achievement in mathematics (Table 7). These factors explained 8.7% of the overall variance in mathematics achievement, as shown in the step-2 model.

In contrast to TIMSS 2011, the gender gap (in favour of boys) was not significant in 2015, after accounting for the power of other variables included in the models (steps 1–3 in Table 7). Similar to TIMSS 2011, books at home, parental education and schools' disadvantaged status were significantly related to student mathematics achievement. In TIMSS 2015, country of birth was also measured. In the third model for TIMSS 2015 (Step 3), the interaction between country of birth and language spoken at home was significant.⁵ The positive sign of the interaction term in the model indicates that students who were born in Ireland and spoke the language of the test at home were expected to perform much better (more than half a standard deviation—29.3 score-points) than other groupings of students—a pattern that was not observed in TIMSS 2011. The step-3 model explained 12.4% of the total variance in mathematics achievement in TIMSS 2015.

Comparing the step-2 models for TIMSS 2011 and 2015, it can be seen that, apart from gender, which was statistically significant only in the TIMSS 2011 model (where boys performed better than girls), the rest of the results were consistent across both cycles. However, the overall variance explained by the examined factors was higher in TIMSS 2015 ($R^2 = 11.9\%$), compared to 2011 ($R^2 = 8.7\%$). This finding suggests that there was no improvement in equalities, but rather there may have been a small increase in inequalities between the two TIMSS cycles, as the examined demographic and socioeconomic factors explained a larger proportion of the variance in mathematics achievement.

To examine whether the changes observed between 2011 and 2015 were statistically significant, TIMSS data from both cycles were included in a single model and the interactions of the year with each one of the examined explanatory variables were tested, after controlling for the main effects (Combined model in Table 7). Even though most

⁵ Significance tests for the main effects (i.e., language spoken at home and country of birth) should be interpreted with caution in the presence of an interaction effect.

interaction terms indicated increased gaps in TIMSS 2015 compared to 2011 (except for gender), none was statistically significant.

Comparing the NAMER and TIMSS multilevel models

Both studies yielded consistent results in terms of the between-school variance in student mathematics achievement, with lower ICCs in the latest cycles of each study; this suggests greater equality across schools. As shown in the step-1 models for each study (which contain variables measured in both TIMSS and NAMER, allowing for comparisons across studies), the results regarding the extent to which the examined demographic and socioeconomic factors explained 4th- and 6th-grade students' mathematics achievement were not always consistent across cycles. Mixed results were found for the role of language spoken at home in explaining mathematics achievement. Boys consistently outperformed girls in all four models and this gap was statistically significant in most cases, even after accounting for other variables in each model. Books at home was the most consistent explanatory variable of mathematics achievement across grade levels and studies, with students whose families had more than 100 books at home performing significantly better than their peers, even after accounting for other factors, including school socioeconomic status. With the exception of NAMER 2014, students attending disadvantaged schools had significantly lower performance than their counterparts in non-disadvantaged schools.

The overall variance in mathematics achievement explained by the examined factors in the step-1 models (i.e., gender, books at home, language spoken at home and schools' disadvantaged status) for both cycles of each study ranged from 5.4 to 10%. This indicates that these factors consistently explain a similar, and relatively small, amount of variance, across studies, grade cohorts and years. However, although the variance explained in NAMER decreased from 10% in 2009 to 5.4% in 2014, suggesting increased equality, this was not the case in TIMSS, where the explained variance increased by 1.6 percentage points.

Discussion

This study conducted an in-depth investigation of changes in educational inequalities in Ireland using national and international assessment data. Drawing on Ferreira and Gignoux's (2014) framework for conceptualising and measuring inequality in education, this paper examined changes in the degree of variability in student mathematics performance (interpreted as an indicator of inequality of achievement) and the extent to which demographic and socioeconomic characteristics explain student mathematics achievement (interpreted as indicators of inequality of opportunity), across different cycles of national and international studies. The study focused on a critical period before and after the introduction and initial implementation of the National Literacy and Numeracy Strategy 2011–2020 in Ireland.

After the introduction of the Strategy, significant improvements were observed in primary students' mathematics performance on NAMER and TIMSS. This study investigated how these changes in performance were distributed across subgroups of students.

The results showed a decrease in the variability in students' mathematics performance in both NAMER and TIMSS, as indicated by the respective standard deviations for each

cycle of each study; this decrease was statistically significant only for TIMSS. According to Ferreira and Gignoux (2014), this provides some evidence of improved equality, since it demonstrates that between-student differences in achievement are getting smaller. International trends in TIMSS data show that, on average, the distribution of mathematics scores in 4th grade has shrunk considerably since 1995, which might indicate a long-term, global reduction in inequalities (Mullis, Martin, & Loveless, 2016). However, during the more recent period from 2011 to 2015, the international average standard deviation did not change significantly; during this same time, the standard deviation for Ireland has notably decreased. This finding provides valuable evidence for Ireland, supporting the argument that equality and efficiency (i.e., high average performance) in education need not be mutually exclusive (Freeman et al., 2010; Mullis, Martin, & Loveless, 2016).

Another piece of evidence that complements this finding is the reduced variance attributed to between-school differences. The between-school variance in mathematics achievement decreased by 36.6% in NAMER 2014 and 42.4% in TIMSS 2015, compared to the previous cycles of these assessments, prior to the introduction of the Strategy.⁶ The reduced ICCs suggest that, after the initial implementation of the Strategy, students' performance in mathematics tended to be less "dependent" on the school they attended. In other words, students had a better chance of performing well, independently of the school they attended. It should be acknowledged that Ireland consistently appears among the countries with the lowest between-school differences in TIMSS, a characteristic that is usually shared by high-achieving countries (Martin et al., 2013).

As far as inequality of opportunity is concerned, this study provided mixed results. In NAMER, all examined groups of students saw improvements in their mathematics performance. These improvements particularly favoured groups of students who had lower performance than their peers in NAMER 2009, and showed increased achievement in NAMER 2014. Improvements in equality were also evident in the results of the multilevel analysis for NAMER, where the variance in mathematics performance attributed to demographic and socioeconomic factors in 2014 was about half as small as it had been in 2009. This decrease can be mainly attributed to a significantly smaller performance gap between boys and girls in 2014 compared to 2009, as indicated by the relevant interaction term.

However, the results for TIMSS do not follow the same pattern. The analysis of 4th-grade students' scores showed that certain performance gaps (e.g., those attributed to educational resources and parent education) increased after the introduction of the Strategy, despite Ireland's overall improved performance. As a result, the variance in 4th grade mathematics achievement explained by the examined factors increased in TIMSS 2015 compared to 2011. In contrast to the NAMER results for 6th-grade students, these results from TIMSS suggest persistent or even increased inequalities, as students' performance seemed to be more "dependent" on

⁶ It is also important to note that, due to the application of two-level analysis, the between-classroom variance in schools with two selected classes is confounded with the between-school variance.

background characteristics; however, it should be acknowledged that no variable significantly contributed to this direction in the multilevel analysis.

The Strategy introduced a number of measures in relation to mathematics including increases in instructional time, a lengthening of teacher education programmes at primary and post-primary levels, additional resources for children in the most disadvantaged areas, additional support for parents, a requirement for schools to submit aggregated standardized test results to the Department of Education and Skills each year, and setting of national targets for the National Assessments. As each of these and other initiatives were implemented to varying degrees, it is difficult to identify which specific initiatives contributed to the changes in equality documented in the current study, and why certain results were not consistent across NAMER and TIMSS. However, since, at primary level, national targets were set only for 2nd and 6th grade, there may have been a stronger focus by schools and teachers on achieving targets at these grade levels. This could partially explain the different findings in relation to equality of opportunity between NAMER (6th grade) and TIMSS (4th grade). It should also be acknowledged that although both NAMER and TIMSS are curriculum-based and have a similar assessment framework, there are differences in the way the two studies conceptualize and measure mathematics achievement.

In the context of this research, inequalities were examined from many different perspectives, something that further enhances the validity of our arguments. However, it is important to acknowledge that the use of only two measurement points for each study used, one before and one after the introduction of the Strategy, provides a limited perspective.

Conclusions

This paper has made several references to changes in performance and equality following the initial implementation of the National Literacy and Numeracy Strategy in Ireland. The gains made in achievement in the National Assessments were also observed in TIMSS 2015, compared with TIMSS 2011. Improvements in a country's overall performance in national and international assessments are always desirable, especially since countries with higher overall performance have been shown to have significantly more resilient⁷ students (Erberber et al., 2015), something that indicates that the effects of socioeconomic inequalities on academic success can be ameliorated. However, as this study suggests, to better examine inequalities, other aspects of student performance (e.g., variability of student scores and variance explained by certain background factors) should be taken into consideration. NAMER and TIMSS yielded consistent results in terms of the variability in 4th- and 6th-grade student performance in mathematics, providing evidence of improved equality in achievement. However, the two studies gave contradictory results in relation to achievement gaps attributed to demographic and socioeconomic factors.

Although decreased standard deviations can be perceived as evidence of improved equality, careful examination of the distribution of the scores across different

⁷ Resilient students are those who, despite their socioeconomic profile, manage to have exceptionally high performance.

performance groups is required to confirm if progress has been made in the desired direction. As Pitsia (2021) confirmed by drawing on national and international assessment results for Ireland, improvements in the country's overall mathematics achievement can be mainly attributed to the improved performance of low-achieving students, while high-achieving students did not improve to the same extent. This may explain, at least in part, the improved mean scores and the reduced variance in mathematics performance found in the present study. If Ireland had managed to achieve its aim of promoting high achievement, as well as improving performance among lower achievers, increased equality attributed to reduced variability in scores, as reported in this study, might not have been found as variance in mathematics achievement would probably have remained stable. This further highlights the importance of monitoring changes in performance from multiple points of view using multiple sources of evidence.

Overall, the evidence provided by this study demonstrates that Ireland has made reasonable progress in addressing inequality. However, there is room for improvement, as a significant proportion of the variance in primary school student mathematics performance is still attributed to demographic and socioeconomic characteristics, especially in TIMSS. The next series of national and international assessments are expected to provide valuable evidence on the extent to which performance gains and changes in equality follow specific patterns and whether it has been possible to build on them.

Abbreviations

OECD: Organisation for Economic Co-operation and Development; UNESCO: United Nations Educational, Scientific and Cultural Organisation; TIMSS: Trends in International Mathematics and Science Study; PISA: Programme for International Student Assessment; PIRLS: Progress in International Reading Literacy Study; NAMER: National Assessments of Mathematics and English Reading; IRT: Item Response Theory; DEIS: Delivering Equality of opportunity in Schools; IBM: International Business Machines Corporation; IDB: International Data Base; IEA: International Association for the Evaluation of Educational Achievement; ICC: Intraclass Correlation; SE: Standard Error.

Acknowledgements

Our thanks to Dr Vasiliki Pitsia for her advice on managing and analysing the NAMER and TIMSS data.

Authors' contributions

All authors made a substantial contribution to the conception, design, statistical analysis and drafting/writing of this research paper. All authors read and approved the final manuscript.

Funding

Not applicable.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 25 September 2020 Accepted: 25 February 2021

Published online: 20 March 2021

References

- Broer, M., Bai, Y., & Fonseca, F. (2019). *Socioeconomic inequality and educational outcomes: evidence from twenty years of TIMSS, IEA research for education*. Springer.
- Checchi, D., Peragine, V., & Serlenga, L. (2010). *Fair and unfair income inequalities in Europe*. IZA Discussion Paper No. 5025. IZA.
- Clerkin, A., Perkins, R., & Cunningham, R. (2016). *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.
- Cosgrove, J., & Cartwright, F. (2014). Changes in achievement on PISA: The case of Ireland and implications for international assessment practice. *Large-Scale Assessments in Education*, 2(2), 1–17. <https://doi.org/10.1186/2196-0739-2-2>

- Cosgrove, J., & Creaven, A. (2013). Understanding achievement in PIRLS and TIMSS 2011. In E. Eivers & A. Clerkin (Eds.), *National Schools, International Contexts: Beyond the PIRLS and TIMSS Test Results* (pp. 201–240). Educational Research Centre.
- Department of Education and Science. (1999). *Primary School Curriculum: Mathematics*. Stationery Office.
- Department of Education and Science. (2005). *DEIS (Delivering Equality Of Opportunity In Schools): An action plan for educational inclusion*.
- 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. (2017b). *Report on the Review of DEIS*. Social Inclusion Unit, Department of Education and Skills.
- Department of Education and Skills. (2017a). *DEIS plan 2017: Delivering equality of opportunity in schools*.
- Educational Research Centre. (2009). *Framework and test specifications for the 2009 National Assessment of Mathematics: Second and sixth classes* (Issue December).
- Eivers, E., & Clerkin, A. (2012). *PIRLS and TIMSS 2011: Technical Report for Ireland*. Educational Research Centre.
- Erberber, E., Stephens, M., Mamedova, S., Ferguson, S., & Kroeger, T. (2015). *Socioeconomically Disadvantaged Students Who are Academically Successful: Examining Academic Resilience Cross-Nationally*. IEA's Policy Brief Series, No. 5. International Association for the Evaluation of Educational Achievement (IEA). <https://doi.org/10.1177/0022146515584605>
- 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>
- Fleurbaey, M., & Peragine, V. (2013). Ex ante versus ex post equality of opportunity. *Economica*, 80(317), 118–130. <https://doi.org/10.1111/j.1468-0335.2012.00941.x>
- Freeman, R., Machin, S., & Viarengo, M. (2010). *Variation in Educational Outcomes and Policies across Countries and of Schools within Countries*. NBER Working Paper No. 16293. <https://doi.org/10.3386/w16293>
- Greaney, V., & Kellaghan, T. (2008). *National Assessments of Educational Achievement: Vol. 1. Assessing National Achievement Levels in Education*. The World Bank.
- IBM. (2017). *IBM SPSS statistics 25 Brief Guide*. IBM Corporation.
- IEA. (2019). *Help Manual for the IEA IDB Analyzer (Version 4.0)*.
- 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>
- Karakolidis, A., Duggan, A., Shiel, G., & Kiniry, J. (2021). Educational inequality in primary schools in Ireland in the early years of the National Literacy and Numeracy Strategy: An analysis of National Assessment data. *Irish Journal of Education*, 44.
- Kavanagh, L., Shiel, G., & Gilleece, L. (2015). *The 2014 National Assessments of English Reading and Mathematics: Vol. 2. Context Report*. Educational Research Centre.
- Kavanagh, L., Weir, S., & Moran, E. (2017). *The Evaluation of DEIS: Monitoring Achievement and Attitudes Among Urban Primary School Pupils from 2007 to 2016*. Educational Research Centre.
- Kellaghan, T., Greaney, V., & Murray, T. S. (2009). *National Assessments of Educational Achievement: Vol. 5. Using the Results of a National Assessment of Educational Achievement*. The World Bank.
- Martin, M. O., Foy, P., Mullis, I., & O'Dwyer, L. (2013). Effective schools in reading, mathematics, and science at the fourth grade. In M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade—implications for early learning* (pp. 109–178). TIMSS & PIRLS International Study Center.
- Mullis, I. V. S., & Martin, M. O. (2013). *TIMSS 2015 Assessment Frameworks*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).
- Mullis, I. V. S., Martin, M. O., & Loveless, T. (2016). *20 Years of TIMSS: International Trends in Mathematics and Science Achievement, Curriculum, and Instruction*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén.
- OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. PISA, OECD Publishing. <https://doi.org/10.1787/9789264073234-en>.
- OECD. (2019). *PISA 2018 results: Vol. 1. What students know and can do (Vol. I)*. PISA, OECD Publishing. <https://doi.org/10.1787/5f07c754-en>.
- Perkins, R., Cosgrove, J., Moran, G., & Shiel, G. (2012). *PISA 2009: Results for Ireland and changes since 2000*. 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 [Unpublished doctoral thesis]*. Dublin City University.
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances* (pp. 91–116). Russell Sage Foundation.
- Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142–151. <https://doi.org/10.3102/0013189X10363170>
- Shiel, G., Kavanagh, L., & Millar, D. (2014). *The 2014 National Assessments of English Reading and Mathematics: Vol. 1. Performance Report*. 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.

- Tarling, R. (2009). *Statistical modelling for social researchers: Principles and practice*. Routledge.
- The World Bank. (2005). *World Development Report 2006: Equity and development*.
- UNESCO. (2018). *Handbook on measuring equity in education*. UNESCO Institute for Statistics.
- United Nations. (2019). *International Migrant Stock 2019 (United Nations database, POP/DB/MIG/Stock/Rev.2019)*. Department of Economics and Social Affairs, Population Division, United Nations.
- Van de Werfhorst, H. G., & Mijs, J. J. B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, 36(1), 407–428. <https://doi.org/10.1146/annurev.soc.012809.102538>
- What Works Clearinghouse. (2020). *What Works Clearinghouse Standards Handbook, Version 4.1*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- 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.
- Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (1996). *BILOG-MG: Multiple-Group IRT Analysis and Test Maintenance for Binary Items [Computer Software]*. Scientific Software International.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
