How do English children fare in international comparisons of mathematical performance?

Sue Pope

University of Manchester

If the rhetoric of politicians is to be believed then the mathematical achievement of English school children is dire and declining or at best stagnating. Close scrutiny of the data reveals a complex picture that is not consistent with this simplistic political message. Drawing on England's outcomes in PISA and TIMSS I attempt to illuminate some of this complexity. International comparisons also reveal interesting differences in practices between primary and secondary phases.

Keywords: international comparisons

Introduction

Considerable changes to the English mathematics curriculum are being introduced, predicated on poor performance in international comparisons (DfE, 2010). Politicians around the world are quick to seize on the outcomes of international comparisons and use them as a basis for changes in education policy without questioning the extent to which the outcomes are valid or reliable (Meyer & Benavot, 2013). A country's position in the list of results is presented as an absolute truth, without any acknowledgement of the associated error bars that indicate clusters of results as opposed to a clear rank order. Although reports carry health warnings around the use of rank orders, because of the clustering as many results do not differ significantly. International comparisons are often presented as highly reliable measures and the differences between the different surveys are neither acknowledged nor explained. In this paper I draw together similarities and differences between PISA and TIMSS, examine trends in performance and other evidence collected as part of each survey.

PISA and TIMSS Similarities and Differences

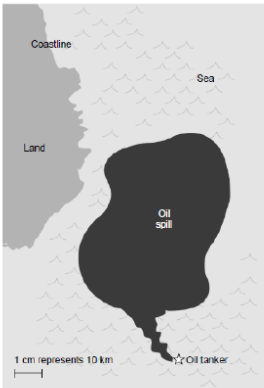
The Programme of International Student Assessment (PISA) is developed by the Organisation for Economic Cooperation and Development (OECD) and is taken by a sample of 15 year olds in participating countries. In the UK, PISA is administered on behalf of governments by the National Foundation for Educational Research (NFER). PISA assesses 'mathematical literacy'

> '…an individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts, and tools to describe, explain, and predict phenomena. It assists individuals in recognising the role that mathematics plays in the world and to make the well-founded judgements and decisions needed by constructive, engaged and reflective citizens.' (OECD, 2013)

Mathematical literacy is probably more closely aligned with functional skills (introduced in England alongside the new secondary curriculum in 2008) than GCSEs. Although GCSE 2010 includes 'functional elements', the proportion is modest (35% at Foundation tier and 25% at Higher tier (Ofqual, 2008). Functional Skills qualifications do not assess mathematical performance per se, rather the candidates' ability to use mathematics to work on situations that an adult might plausibly meet in their everyday life (Ofqual, 2009). One of the potential issues with PISA items is the literacy demands.

OIL SPILL

An oil tanker at sea struck a rock, making a hole in the oil storage tanks. The tanker was about 65 km from land. After a number of days the oil had spread, as shown on the map below.



Using the map scale, estimate the area of the oil spill in square kilometres (km^2).

Answer: .... km^2

Peter, James, and Andrew each had 20 tries at throwing balls into a basket.

Complete the missing boxes below.

Name	Number of Successful Shots	Percentage of Successful Shots
Peter	10 out of 20	50 %
James	15 out of 20	$\boxed{75}$ %
Andrew	$\boxed{16}$ out of 20	80%

Figure 1: PISA item (Wheater et al., 2013: 31) and TIMSS Y9 item (Sturman et al., 2012:217)

The Trends in International Mathematics and Science Study (TIMSS) is developed by the International Association for the Evaluation of Educational Achievement (IEA) and is taken by a sample of 9-10 year olds and 13-14 year olds every four years in participating countries. In the UK the tests are administered by NFER on behalf of governments. TIMSS items assess those elements of the school curriculum that are common across participating countries and include items that teachers are likely to be familiar with, including arithmetic and algebra.

The sample items (Figure 1) illustrate the difference between the assessments – the PISA item expects students to be able to develop a strategy for dealing with an unstructured problem, whereas the TIMSS item is a straightforward assessment of the use of percentages to describe proportions. Contextual data is also collected including the amount of time spent teaching the subject, student attitudes and dispositions and aspects of teachers' professional life.

PISA outcomes

In 2012 PISA included a computer delivered assessment of problem solving skills. English children did well in this assessment (see figure VA in OECD, 2014: 15). PISA scores are scaled so that the average is 500 with a standard deviation of 100. In this assessment involving 44 countries with average scores ranging from 399 to 562, the results cluster with groups of countries whose scores are not statistically significantly different.

England's results were not statistically significantly different from Canada, Australia or Finland who formed a cluster after the typically high performing Pacific Rim countries and the cluster below which included the US and some European countries.

Trends over time in PISA mathematics assessments

Performance in PISA has been popular in government rhetoric around the decline in English academic standards (DfE, 2010 Para 4.36). The position in the PISA league tables has been cited using assessment in 2000 as a baseline, despite the fact that England's sample meant that the results were not valid and an increasing number of countries participate in PISA.

As shown in Table 1, England's mathematics scores in PISA 2006, 2009 and 2012 have remained similar to the OECD average. The difference in performance between Wales and the rest of the UK is statistically significantly different. PISA 2012 focussed on mathematics and revealed relative strengths in probability and interpretation, and relative weaknesses in geometry and modelling. The opposite trends were noted in high performing Pacific Rim countries (Wheater et al., 2013).

	2006	2009	2012
England*	495 (2.5)	493 (2.9)	495 (3.9)
Northern Ireland*	494 (2.8)	492 (3.1)	487 (3.1)
Scotland	506 (3.6)	499 (3.3)	498 (2.6)
Wales*	484 (2.9)	472 (3.0)	468(2.2)
United Kingdom	495 (2.1)	492 (2.4)	494 (3.3)
OECD average	498 (0.5)	496 (0.5)	494 (0.5)

Table 1 England's results (standard error in brackets) in PISA tests 2006, 2009 and 2012

In 2012, seven of the countries that significantly outperformed England were EU members (Austria, Belgium, Estonia, Finland, Germany, Netherlands and Poland). A further eight EU countries did not perform significantly differently from England (Czech Republic, Denmark, France, Latvia, Luxembourg, Portugal, Republic of Ireland and Slovenia) and 11 performed less well (including Hungary, Italy, Spain and Sweden). Among OECD countries, 12 outperformed England (including Australia, Canada, Japan and Switzerland), 10 performed similarly (including Iceland, New Zealand and Norway), and 11 performed less well (including Israel and the US). This indicates that England, while not among the highest achieving group of countries internationally, compares well with other EU and OECD countries in terms of mathematics achievement.

Boys out perform girls in most countries in PISA although the gap is greater in England, Scotland and Wales. The gap is negligible in Northern Ireland.

Trends over time in TIMSS outcomes

The TIMSS survey allows comparison of performance of a sample from the same cohort at age 9-10 and then at age 13-14. The table below summarises England's average performance in TIMSS since 1995. In contrast to PISA there are no gender differences in performance on TIMSS.

	1995	1999	2003	2007	2011
England Age 9-10*	484 (3.3)	-	531 (3.7)	541 (2.9)	542 (3.5)
England Age 13-14	498 (3.0)	496 (4.1)	498 (4.7)	513 (4.8)	507 (5.5)

Table 2 England's mathematics performance (standard error in brackets) in TIMSS 1995 -2011

In TIMSS 2011 England's 9-10 year olds' performance was significantly higher than the average, whereas 13-14 year olds' performance was close to average. This was true for a number of other countries including Australia, Hungary, Italy and

Lithuania. In 2007, England's performance at both ages was significantly above average (Sturman et al., 2008). In 2011 only Pacific Rim countries performed more highly than England. Figure 2 shows how England's mathematics performance (dark rectangular outline) is in a similar relative position at both age 9-10 and age 13-14 to the top of the list but have a very different relationship to the TIMSS average (at the bottom of each list)

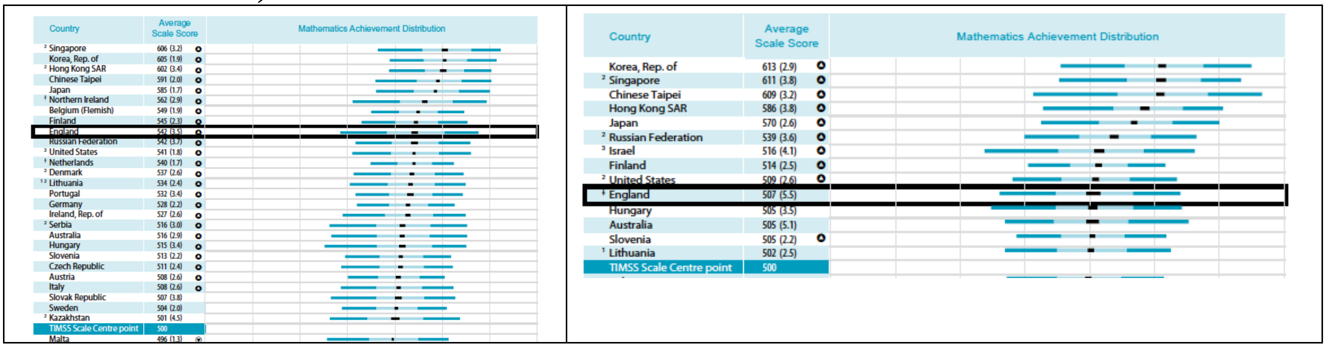


Figure 2 Extracts from TIMSS 2011 tables for primary and secondary (Sturman et al., 2012: 4 and 8)

The content analyses suggest that children's strengths are with items that assess data, which is also found in PISA. This may be attributable to the fact that statistics and probability are a substantial part of the National Curriculum for mathematics since its inception in 1988.

Other contextual factors

TIMSS includes a survey of students, teachers and headteachers to give rich data about the school context. As with any survey the validity of the responses is dependent on the quality of the instrument (Oppenheim, 1992). Table 3 compares English averages with the TIMSS average for selected contextual factors. These have been selected because of the substantial difference from the TIMSS (International) average or between the different phases.

	Source	Age 9-10	Age 13-14
Text book use (Int. average)	TiMSS 2007	15% (3.1) 65% (0.5)	43% (4.1) 60% (0.5)
Teaching time (Int. average)	TIMSS 2011	188 (3.3) hrs p.a. 162 (0.5) hrs p.a.	116 (2.1) hrs p.a. 138 (0.5) hrs p.a.
Use of computers (Int. average)	TIMSS 2011	71% (4.2) 42% (0.5)	51% (4.3) 36% (0.5)
Specialist teacher (Int. average)	TIMSS 2011	19% (3.1) 38% (0.5)	76% (4.0) 73% (0.5)
Teacher Collaboration (Int. average)	TIMSS 2011	47% (4.0) 36% (0.5)	24% (3.8) 28% (0.5)

Table 3 Comparison of selected contextual factors for England with the TIMSS average

The relatively poor quality of English text books has been established elsewhere (Haggarty & Peppin, 2002). It is striking how little use is made of textbooks in English schools, particularly in the primary phase. Across all TIMSS participants the difference in teaching time at age 9-10 and age 13-14 is equivalent to less than one hour per week, whereas in England the difference is two hours per week (most primary schools have a daily one hour mathematics lesson whereas secondary school provision varies considerably but three one hour lessons is commonplace). These figures have informed the government recommendation that schools will provide an additional hour of teaching each week for 14-16 year olds for the new GCSE from 2015 (DfE, 2013).

Teachers in England have far greater access to computers compared with other TIMSS participants. However, it is not clear what the computers are used for, indeed there is evidence that the increased use of interactive whiteboards has decreased the opportunity for hands on use of technology for individual exploration of mathematics (Ofsted, 2008).

Whilst the proportion of specialist teachers at age 9-10 was half that of the international average, over 90% of these teachers felt well prepared to teach mathematics. At age 13-14 the proportion of specialist teachers was not statistically different to the international average and 94% felt well prepared to teach mathematics. This is in stark contrast to the research evidence (Ma, 1999) and recent Ofsted (2008, 2012) reports where weaknesses in teacher subject knowledge are frequently identified as a cause for concern.

The extent to which teachers collaborate differs significantly between the phases. Teachers of 9-10 year olds are far more likely to collaborate to enhance their teaching provision. Teachers of children aged 9-10 in England are significantly more likely to do this, whereas teachers of 13-14 year olds in England collaborate to the same extent as other TIMSS participants, on average.

Attitudes, dispositions and the value of mathematics

Both PISA and TIMSS survey students about their attitudes towards mathematics. All the TIMSS studies have found that the younger cohort like mathematics (44% in 2011) more than the older students (14% in 2011), and that the younger children feel more confident about mathematics (33% compared with 16%). The English average for the younger cohort is similar to the international average for both liking and confidence (48% and 34% respectively), whilst for the older cohort it is much less for liking and similar for confidence (28% and 14% respectively). Between 1995 and 2007 there was a steady decline in the proportions that reported liking mathematics. As discussed by Ruthven (2011) the significant decrease in English children liking mathematics was accompanied by significant increases in performance. In PISA studies 15 year olds liking of mathematics is similar to the international average (approximately half).

TIMSS surveys 13-14 year olds about the extent to which they value the subject and found a higher proportion (48% in 2011) than those who liked the subject and/or felt confident about it, this was similar to the international average. In the most recent PISA study the proportion valuing mathematics was greater than the international average (over 80% compared with 70%).

Conclusion

English children have improved performance in the TIMSS surveys since 1995, particularly at primary, but that improvement has not been sustained between 2007 and 2011. The improvements in primary attainment have not been sustained in secondary, where the improvements are comparatively modest. In recent PISA and TIMSS studies performance of older students (15 year olds and 13-14 year olds) is close to the international average and relatively strong compared with the rest of the Europe.

The TIMSS studies suggest that improved attainment has been at the cost of positive attitudes towards mathematics. Despite this, in both PISA and TIMSS it appears that English secondary students value mathematics highly.

The contextual information gathered as part of these international studies suggests potentially valuable starting points for further exploration including the use of textbooks, digital technology, the amount of time devoted to teaching mathematics, the extent to which teachers are prepared to teach the curriculum and the extent to which teachers collaborate.

References

DfE (2010) *The importance of teaching* https://www.gov.uk/government/publications/the-importance-of-teaching-the-schools-white-paper-2010

DfE (2013) *Reformed GCSEs in English and mathematics written statement to parliament* https://www.gov.uk/government/speeches/reformed-gcses-in-english-and-mathematics

Haggarty, L.& Peppin, B. (2002) An investigation of mathematics textbooks and their use in English, French and German Classrooms: who gets an opportunity to learn what?' , *British Educational Research Journal* 28 (4): 567-90

Ma, L. (1999) *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

Meyer, H.-D. & Benavot, A. (2013) *PISA, power and policy the emergence of global educational governance* Oxford: Symposium Books

OECD (2013) PISA 2012 mathematics framework. In: *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy.* Paris: OECD Publishing [online]. Available: http://www.oecd.org/pisa/pisaproducts/PISA%202012%20framework%20e-book_final.pdf

OECD (2014) *PISA 2012 Results: Creative Problem Solving Students' skills in tackling real-life problems Vol. 5* http://www.oecd.org/pisa/keyfindings/PISA-2012-results-volume-V.pdf

Ofqual (2008) *GCSE 2010 Mathematics criteria* http://ofqual.gov.uk/documents/gcse-subject-criteria-for-mathematics/

Ofqual (2009) *Functional skills Mathematics criteria* http://ofqual.gov.uk/documents/functional-skills-criteria-for-mathematics/

Ofsted (2008) *Mathematics: Understanding the Score* London: HMSO

Ofsted (2012) *Mathematics: made to measure*, London: HMSO

Oppenheim, A. N. (1992) *Questionnaire design, interviewing and attitude measurement*. London: Pinter

Ruthven, K. (2011) Using international study series and meta-analytic research syntheses to scope pedagogical development aimed at improving student attitude and achievement in school mathematics and science. *International Journal of Science and Mathematics Education*, 9(2), 419-458

Sturman, L., Burge, B., Cook, R. & Weaving, H. (2012) *TIMSS 2011: Mathematics and Science Achievement in England* Slough: NFER

Sturman, L., Ruddock, G., Burge, B., Styles, B., Lin, Y. & Vappula, H. (2008) *England's Achievement in TIMSS 2007 National Report for England* Slough: NFER.

Wheater, R., Ager, R., Burge, B. & Sizmur, J. (2013) *Achievement of 15-Year-Olds in England: PISA 2012 National Report* (OECD Programme for International Student Assessment) London: DfE (RR307)