Mid-secondary science in England in TIMSS 2023: what can we learn for the classroom?

Jennie Golding, University College London IOE

Abstract

TIMSS is a large-scale four-yearly international assessment of mathematics and science performance, attitudes and reported experiences. TIMSS assesses two cohorts: 9/10 year olds (in England, year 5) and 13/14 year olds (year 9). In this article I point to key findings around England's year 9 science in the latest cycle of TIMSS, including differential findings by socio-economic, gender, and first language student groups. The focus of the paper, though, is on ways that those close to school science education and teacher education can make productive use of TIMSS findings and resources. I identify at a high level some of the approaches known to differentially positively impact the currently under-performing groups, and show how TIMSS questionnaire and released performance items can be directly harnessed in ways consistent with those aspirations.

What is TIMSS?

TIMSS ('Trends in mathematics and science study') is a four-yearly large-scale international study of (in the case of England) year 5 and year 9 students' performance in science, and assessment of the related attitudes and learning experiences as reported by sample learners and, for the latter, their teachers and Headteachers. For year 9, England has participated in every cycle since 1995, and although samples from the early cycles are not directly comparable with later ones, since 2003 the approach has been continuous, and the use of 'trend items' across two or three successive cycles ensures that patterns in performance over time can be tracked. Further, since all students in the older age group ('grade 8') study are of comparable age (13-14 years), TIMSS data can also be used to analyse international comparative data in relation to the TIMSS curriculum 'framework'. In 2023, 44 mostly medium- and high-income countries participated in the grade 8 study, and in England the sample included 4239 year 9 students, their mathematics and science teachers, and their Headteachers. In 2023, for the first time, all assessment was carried out on-screen.

TIMSS is run by IEA, an international education consortium. In England for TIMSS 2023, DfE commissioned Pearson to lead all aspects of assessment, recruitment and delivery up to the stage of acquiring all raw data. After IEA checking of those data, we at UCL then led on all analysis and reporting of that data; the two parts of the national report, focusing on attainment and attitudes/experiences respectively, can be found here. Reports from other participating countries are also usually in the public domain, and readers based elsewhere can use those to adapt to their own context the classroom-focused applications suggested in this article.

Unlike the OECD's PISA, which measures 15 year olds' scientific literacy, TIMSS assesses students across areas of Biology, Chemistry, Earth Science and Physics that are fairly well-matched to the Key Stage 3 national curriculum in England - although our young and people meet much of the material in 'Earth Science' in their geography, rather than their science, lessons. Assessment items are developed to also target one of core knowledge, reasoning or application in relation to that content. When making international comparisons in relation to attainment, it is important to bear in mind that while the intended curricula in most participating countries have much in common, they do differ in detail; nevertheless, such international comparisons serve at the last to raise questions about the relationship of student attainment in science to education and broader societal systems and contexts.

Comparing international reports of attitudinal and experiential data is more problematic: most such data derive from Likert-scale responses to a 'basket' or related statements, many of which are contextually- or culturally-subjective (for example, notions of 'in-class disruption' vary between cultures), although questionnaire statements are designed to be as objective as possible. It is also centrally important to remember that where analysis shows an association between performance and specific attitudes or experiences, TIMSS does not offer the data needed to definitively establish whether those associations are causal – or if so, in which direction that causality might function.

While longitudinal and comparative data are clearly of interest to policymakers, in this article I focus on aspects of the TIMSS 2023 attainment and questionnaire data that are particularly pertinent to teaching and learning in the classroom – and to teacher education. For released both attainment and questionnaire items, teachers can validly compare their own class's responses with those shown nationally, since our analysis in England shows the national sample is robust in a number of important ways. For the interested, other aspects of the international data can be explored here, and also contextualised in an understanding of the relevant national education system via the high level outlines available in the TIMSS 2023 Encyclopedia.

TIMSS scale scores

TIMSS attainment data are reported on a scale established in 1995, with a 'centrepoint' of 500 (the mean score in 1995, when the SD was scaled to 100). On that same scale, and to support score interpretation, TIMSS identifies achievement at four points along the scale as International Benchmarks: Advanced International Benchmark (625), High International Benchmark (550), Intermediate International Benchmark (475), and Low International Benchmark (400). The 2023 descriptions of science performance at the International Benchmarks were updated from TIMSS 2019 based on an analysis of the items that students with average attainment at each of the benchmarks answered successfully in TIMSS 2023, but broadly speaking, because of the methodology used, international benchmarks are comparable across different series of TIMSS.

How did England's year 9 students perform in science in TIMSS2023?

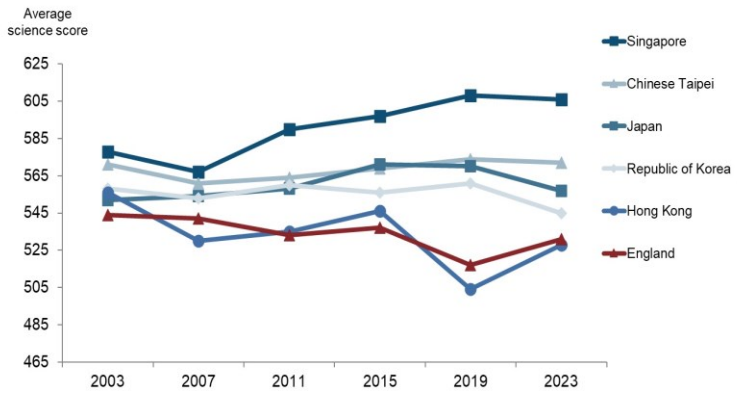


Figure 1: England's year 9 performance compared with TIMSS 2023 highest performing countries, over time

In 2023 England's year 9 average score recovered to near its previous levels after an unexpected dip in 2019, exceeded only by average scores in Singapore, Chinese Taipei, Japan and the Republic of Korea (Figure 1). Given the students in this cohort were in years 6 and 7 in 2020 and 2021, so significantly affected by pandemic-related school closures and the evidenced related sustained impact on learning (e.g. Ofsted, 2023), this finding is welcome news indeed, though it should still be treated with caution. This cohort did not complete national Key Stage 2 assessments, so the representativeness of the sample cannot be checked in as much detail as usual. The 2023 data do suggest, though, that **students' core science curriculum knowledge is, on average, recovering from the pandemic.** Our further analysis compared year 9 average performance in each of the component subject domains, as well as each of the targeted cognitive domains (knowing, applying and reasoning in science) with overall performance, and found there were no significant differences – so **on average, our year 9 students are acquiring scientific knowledge, skills and processes that are broad and balanced.**

That is not to say, however, that there are no causes for concern around year 9 students' performance in science. Chief among these were a **wider range of attainment**, driven by stronger performance at the top end and with stagnation at the weaker end, strongly associated with socio-economic status: significantly smaller percentages of year 9 pupils who had been eligible for free school meals (FSM) in the previous six years reached each of the benchmarks in science compared with their non-FSM eligible peers. TIMSS actually uses a different measure of socio-economic status, namely the number of books students estimate they have in their home. Interestingly in a digital age, that still shows a significant association with performance. Socio-economic status has a persistent, and relatively intractable, relationship with science performance across much of the globe, and one that is clearly established by year 5 (Richardson et al., 2015), though as below, there is much we can do in the classroom to begin to address that.

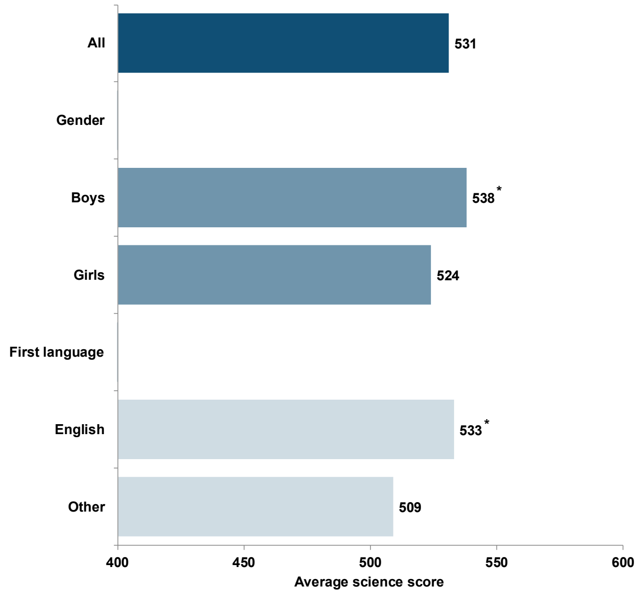


Figure 2: England year 9 average science performance by gender and first language

TIMSS 2023 also showed a **significant gap in average science performance by gender**, in favour of boys (at each benchmark level, and one of the biggest gaps internationally). Figure 2 shows there was also a significant gap in average performance depending on whether English was or was not the **first language of the student**: year 9 students' access to the science curriculum, and to the means to demonstrate their knowledge and skills, can be limited by their language skills.

Addressing demographic attainment gaps in the classroom

These are findings that alert teachers to possible inequalities in curriculum mastery in their own classrooms, but of course they are based on national data, so will not necessarily transfer in any detail. They do, however, raise teacher awareness of possible inequities in access by student demographic. Fortunately, we also have a significant body of work that suggests that good quality science teaching, while supportive of all students, is likely to be differentially supportive for relatively underperforming groups of students. For example, the Improving Secondary Science materials (EEF, 2022) evidence the importance of **starting from students' current knowledge base** (and creatively exposing and addressing existing misconceptions, which we know is likely to be differentially beneficial for less well-resourced students; for **building metacognition** (the ability to analyse, monitor and control one's own learning) and **promoting metacognitive talk and dialogue in the classroom** - differentially important for girls; for **careful attention to development of scientific vocabulary and support for students to read and write about science** – differentially important to students learning science through English as an additional language.

How can year 9 science context, attitude and aspirations scales inform classroom understanding?

The other valuable source of TIMSS data is the range of questionnaire items and responses, in relation to home and school contexts and student attitudes. Attitudes and aspirations are associated with continued and effective participation in science (Archer and DeWitt, 2016; Mao et al., 2021), and a range of evidence suggests performance, attitudes, aspirations and participation are all interlinked in complex ways (e.g. Gorard et al., 2012). Nationally, there are again concerns about the contrast in average responses across many items around attitude, aspirations, and perceived learning contexts, **by both socio-economic status and by gender.**

For example, in relation to gender and the measures used **for confidence in science, for valuing science, and for liking science, boys responded significantly more positively and girls significantly more negatively.** In relation to aspirations – for continuing to study science after age 16, and for doing a job that involved science - there was no significant difference by gender among the 18%/19% respectively who 'strongly agreed', although girls were significantly more likely to 'strongly disagree'.

Both performance and questionnaire items have been very thoroughly developed and extensively trialled before use, so represent robust resources that can very easily be adopted for class use. The internationally-common questionnaire items are all available here, but for England's year 9 students they were supplemented by England-only questions around science aspirations and about practical work in science. Those are reported in detail in sections 8.7/8.8 and section 10.12 of the England Report, respectively, and teachers are able to use those questions with their own students. As always, responses are as useful for the questions they raise as for the specific data they show.

How can classroom teachers support positive attitudes and aspirations, across groups of students?

As suggested above, these positive affective traits interact with learning in complex ways; both are supported by regular attendance, in-class approaches that support development of scientific identity, interest and self-efficacy, and a sense of belonging (De Witt, Archer and Mau, 2016), interacting in a positive cycle. Classroom teaching needs approaches that affirm the student as a scientific thinker and do-er, able to make sense of and interrelate scientific knowledge and skills, and to apply those to new scientific situations. EEF (2022) cites evidence that key approaches to doing so, in addition to those identified earlier, include:

- **Building on the ideas that students bring to class**
- **Using models to support understanding**
- **Using practical work purposefully and as part of a learning sequence**
- **Using structured feedback and focused interactive talk to move on students' scientific thinking**
- **Supporting students to retain and retrieve scientific knowledge**

How can assessment items be harnessed for classroom use?

Finally, TIMSS items and the related performance data are a rich resource for in-class tasks that probe student grasp of key science knowledge, and its associated reasoning and application. They also offer opportunity well-trialled opportunities for retrieval of knowledge, and for making links across different areas of science. Many TIMSS items are not released, in order to support comparability across successive series, but the TIMSS website includes released items tied to most benchmarks, together with related international performance data, for most cycles of TIMSS: for example, science items released for TIMSS 2019 can be found here. The performance data support teachers in understanding what it is that students find most difficult – and how England's average, or indeed, their own class average, performance, compares with international norms. Released items in any one series span each content and each cognitive area, as exemplified below. Readers might like to consider, in each case shown, where the likely learning challenges are, and what pedagogical approaches the teacher might take to support exploring the related ideas and to 'bridge the gaps' exposed. How might such an item, or others available on the TIMSS website, be harnessed to enrich learning in a science classroom – or, indeed, when working with beginner science teachers? Making full formative use of student (or teacher) responses, in ways that align with the science pedagogical principles suggested above, is, of course, demanding on both subject knowledge and subject teaching skill. Each item shown includes average facility, that is, the percentage of sample students gaining full marks on that item.

Example 1: Intermediate benchmark Biology reasoning item (International average 55%; England 67%, significantly higher)

Dixon read a fact sheet about crocodiles.

Crocodile Facts

1. Crocodiles have a lifespan of up to 75 years.
2. Crocodiles today look like ancient crocodiles found in fossils.
3. Crocodiles have an angle of vision of 290° as shown in the diagram.

How can a crocodile's angle of vision help it to survive in its environment?

Give one reason.

Example 2: High benchmark Physics reasoning item (International average 38%, England average facility 35%, not significantly different)

Nada hangs her cell phone under a glass bowl as shown. The ringer on the phone is turned on. She removes the air from under the bowl so that her phone is in a vacuum.

Nada asks her friend to call her phone. Will they hear it ring?

(Click one box.)

☐ Yes

☐ No

Explain your answer.

Example 3: Advanced benchmark Chemistry application item (International average 29%, England average facility 31%, not significantly different)

This is a portion of the periodic table of elements.

1 H							He
Li	Be	B	C	N	O	F	Ne
Na	Mg	Al	Si	P	S	Cl	Ar

Hydrogen (H) is the first element of the periodic table. The nucleus of a hydrogen atom contains one proton. The atomic number of hydrogen is 1.

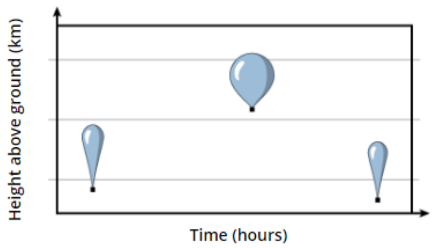
Four elements from the periodic table are shown below. The elements are not ordered by their atomic numbers.

Drag the four elements below to sort them by atomic number from smallest to largest.

Smallest **Largest**

Example 4: Advanced benchmark Earth Science knowledge item (International average 42%, England average facility 42%, not significantly different)

The diagram shows the height above the ground of a helium-filled weather balloon during a period of several hours.



What causes the balloon to become bigger as its height above the ground increases?

A Gravity decreases.

B Atmospheric pressure decreases.

C The balloon is heated by the Sun.

D The balloon absorbs air.

Conclusion

Analysis of England's TIMSS 2023 science data therefore exposed important findings, nationally: on average, year 9 students appear to have largely recovered pre-pandemic levels of scientific knowledge and skills – and in a broad and balanced way across the science curriculum. However, socio-economic gaps in performance have widened further since previous (pre-pandemic) cycles of TIMSS, performance gaps by gender in 2023 widened considerably, and we are reminded that students' grasp of English is critical for their thriving in science. There are, though, evidence-based approaches to at least shrinking some of those gaps, in ways that are likely to benefit all students, but to differentially impact under-performing groups.

Although national findings are of interest to science teachers and teacher educators, not least since they are likely to influence education policy, available TIMSS materials also offer opportunities to explore particular science knowledge and skills – and importantly also, science aspirations, attitudes and perceptions of science experiences, at a more local level, and to contextualise those within a broader national picture. In particular, released TIMSS items offer opportunities to apply the identified productive pedagogical approaches across the science curriculum.

References

Archer, L., & DeWitt, J. (2016). *Understanding Young People's Science Aspirations: How students form ideas about 'becoming a scientist'*. Routledge.

DeWitt, J., Archer, L. and Mau, A. (2016). 'Dimensions of science capital: exploring its potential for understanding students' science participation'. International Journal of Science Education, 38 (16), 2431-2449.

Education Endowment Foundation (2022). *Improving secondary science.* Improving Secondary Science | EEF

Gorard, S., See, B. H., & Davies, P. (2012). The impact of attitudes and aspirations on educational attainment and participation. *York: Joseph Rowntree Foundation*.

Mao P, Cai Z, He J, Chen X and Fan X (2021). *The Relationship Between Attitude Toward Science and Academic Achievement in Science: A Three-Level Meta-Analysis*. Front. Psychol. 12:784068. doi: 10.3389/fpsyg.2021.784068

Ofsted (2023). Finding the optimum: the science subject report. Finding the optimum: the science subject report - GOV.UK

Richardson M., Golding, J., Isaacs, T., Barnes, I., Swensson, C., Wilkinson, D. (2024). *Trends in International Maths and Science Study (TIMSS) 2023 National Report for England Volume 2*. London: Department for Education. https://www.gov.uk/government/publications/trends-in-international-mathematics-and-science-study-2023-england

Jennie Golding is Associate professor and PI, TIMSS 2023 England analysis and reporting, University College London.