
How can we measure educational success?

Tanya Wilson
University of Glasgow, IZA

20th May 2024

Who are the key “stakeholders”?



Measurement Methods

- Standardised Testing - measure students' proficiency in specific subjects.
- Non-standardised testing - class/school based assessments
- Teacher Assessments (of students) - eg reports reflecting overall achievement. (Or potential?)
- Longitudinal Assessments - Tracking students' progress over time to assess long-term educational outcomes
- Entrance exams - eg SAT scores in US
- Assessments of teachers - peer observation, student performance, inspections/audits
- School Inspections and Audit - evaluate the quality of education provided by schools, including effectiveness of teaching, management etc.
- Participation and Engagement Metrics: e.g. measuring students' attendance, participation in extracurricular activities
- Parent and Community Feedback - e.g. perceptions of school's effectiveness

- Subject Specific
 - Scotland: Nationals/Highers
 - Rest-UK: GCSE/Alevel
 - France: Brevet, Le Bac
- Subject Range
 - Germany - Abitur
 - Israel - Bagrut

US– HS diploma is not a national standardised test. “Passing criteria” decided locally (sometimes school specific)

- Additional testing (SAT) used to assess “readiness” for HE

Advantages/Disadvantages of standardised testing

Pros:

- Performance metrics can motivate students
- Allows comparison of performance across different groups of students *
- Allows comparison of performance across different groups of schools
- Can serve as a measure of teacher effectiveness/identify areas where students excel or need additional support
- Consistent testing over time provides valuable data on educational trends, helping policymakers and educators make informed decisions. *

* - requires access to *individual student data* for analysis

Advantages/Disadvantages of standardised testing

Cons:

- The high stakes associated with standardized tests can cause significant stress and anxiety for students, potentially affecting their performance and well-being. (Teachers and parents too)
- Teachers may feel pressured to limit their curriculum to tested subjects
- Can lead to a narrow focus on test preparation, neglecting broader educational goals
- Can amplify SES disparities - eg test preparation resources (incl tutoring), environments
- One size does not always fit all
- Administration cost (including time cost)
- Ability to make cross-country comparisons may be limited - eg GCSE vs National 5
- Overemphasis on aggregate school metrics, not value-added

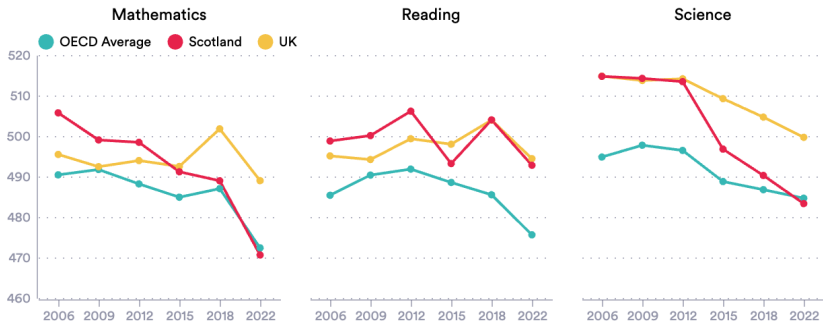
Standardised Tests:

- Programme for International Student Assessment (PISA)
 - Reading/Mathematical/Scientific Literacies
 - Student Age: 15
 - 80+ countries
 - Every 3 years
- TIMSS (Trends in International Mathematics and Science Study)
 - Mathematics/Science
 - Student Age: 9/10 and 13/14
 - 60+ countries
 - Every 4 years
- PIRLS (Progress in International Reading Literacy Study)
 - Reading Comprehension
 - Student Age: 9/10
 - 60+ countries
 - Every 5 years

PISA performance

Student Performance

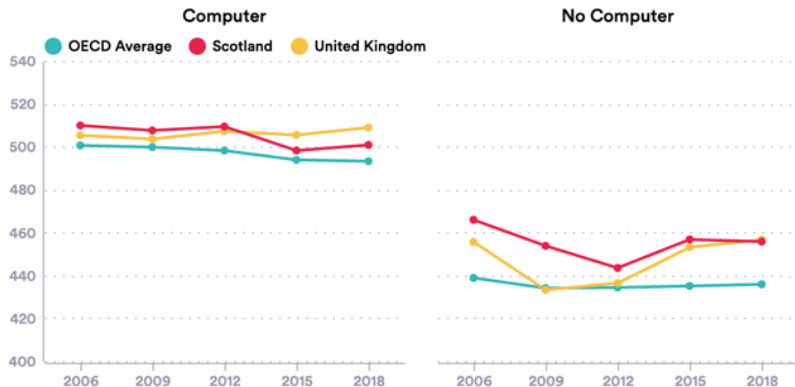
Source: PISA Test Scores (OECD)



PISA performance

Student Performance by Computer Access

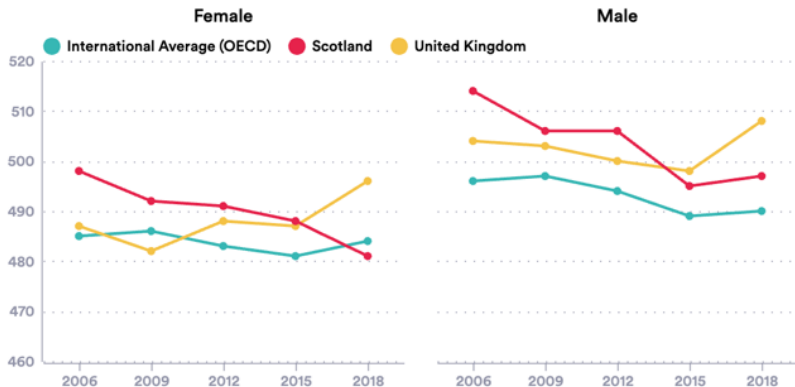
Source: PISA Test Scores (OECD)



PISA performance

Student Performance by Gender

Source: PISA Test Scores (OECD)



- In-class/school metrics - quizzes/projects/homework
- Teacher-based assessments/judgements used during Covid to ensure students received qualifications
 - Teacher judgements based on a range of evidence collected throughout the academic year.
 - Used various types of evidence to form a holistic judgement of a student's abilities and achievements.
 - Internal and external moderation

Advantages/Disadvantages of non-standardised testing

Pros:

- Evidence reflects on-going performance rather than a one-off exam.
- Absence of high-stakes exams reduces stress and anxiety for students
- Potentially allows for more equitable assessments

Cons:

- Challenges to ensure consistency and standardisation across different schools and teachers
- Potential unconscious bias
- Significant increase in workload for teachers

Wyness et al (2023) use a statistical model inputting previous achievement (GCSEs) to predict future performance (A Levels), finding it worked better than teacher predictions. (27% vs 16%)

Is test data everything?

Tests are particularly useful at measuring things “easy” to test

Test-score data tells us “how much/how many”, but not necessarily the “why”

Linked data: test scores + student characteristics + school characteristics allows an assessment of the “why” and also future consequences

Data on other educational outcomes and qualitative data also help with the “why” and “what works”.

- Authorised/unauthorised absence (participation in education)
- Opinions (students, teachers, parents)
- Confidence and other non-cognitive skills

Case Study: South Korea

Education System Overview:

- Intense focus on education
- High PISA scores

Key Metrics:

- High-stakes exams
- Competitive school environment
- Hagwons (Private Tutoring Centers) used by many to supplement education - high cost, increased academic pressure, can exacerbate educational inequalities

Pros:

- High academic achievement
- Strong societal value on education

Cons:

- High student stress
- Limited creativity

- **Key Stakeholders:** Many “users” of education outcomes data
- **Measurement Methods:** Variety of ways to measure outcomes, each have their pros and cons
 - **Standardised Tests:** Useful to be able to compare across different groups and over time, high administrative (and cost) burden
 - **Non-standardised Testing:** Potentially more equitable assessment, but additional workload burden for teachers
 - **Non-test data:** Additional educational outcomes and qualitative data provide for a comprehensive understanding of educational effectiveness.
- **Research Impact:** Ability for research to inform the evidence base depends both on quality and availability of data.