

Pearson World Class Qualification

International Comparative Analysis of Science and Maths Assessments

Contents

The issue.....	2
What does the research tell us?.....	2
In summary	3
Next steps.....	3

The issue

International benchmarking has become an important tool for politicians and others seeking to raise standards in the education system. The present administration, for instance, has placed a lot of emphasis on lessons to be learned from high-ranking countries in PISA, TIMSS and other international tests.

However, others have argued that such tests provide only a partial picture of how a particular education system is performing, that education systems and priorities differ from country to country, making meaningful comparisons difficult and that what works in one country may not easily be replicated in another.

A key issue is how different countries measure their own performance, which assessment instruments they use and how valid their performance judgements are perceived to be. It has been argued, for example, that high-performing PISA countries adopt a more rigorous approach to the prescription and assessment of knowledge in core subject areas such as Maths and Science. The Secretary of State believes this is so and has called for a similar approach to be adopted for reforms in England.

What does the research tell us?

This is an area that has attracted considerable research.

This particular piece of research examines the assessment instruments used by five anonymised, high-performing countries in the core subject areas of Maths and Science (including Biology, Chemistry, Physics and Science) and compares them with those used in England to see what, if any, conclusions can be drawn.

Five critical features have been used in determining conclusions:

- **content mapping** against generic international standards;
- the **cognitive operation** for each question item, in effect the manner in which items are presented and the intended nature of response;
- **cognitive demand**, such as task complexity in assessment items;
- **question type** such as open-ended or extended response;
- the **number of marks** per item.

Overall conclusions:

- **Appropriate content representation** is a key assessment issue. This should be monitored longitudinally and incorporated into the design principles for any new qualification adopted.
- Many systems demonstrated a **predominance of memory and recall items**, but how this was approached varied considerably.
- **Higher-order thinking skills** were not well represented in some assessments. This was particularly evident for problem-solving skills in Maths and for applying concepts and making connections in Science.
- **Cognitive operations** and **cognitive demand** often work independently of each other and therefore need to be actively designed into assessments.
- Assessed knowledge should consist of a **range** of cognitive operations at a range of cognitive demands and this needs to be actively planned for rather than assumed.

- **A range of different question types** was evident. The analysis showed that although more closed question types can assess a range of cognitive operations and demands, they are less effective than more open responses.
- Although the criticism of '**mile wide and inch deep**' assessment in Science and Maths remains, this particular research found little evidence of this operating in practice.
- The **number of calculation items** found in science assessments were relatively low, especially in Biology and Chemistry. This supports the view that there is a general lack of mathematical skills in science assessments.

In summary

Two fundamental points emerge from this research.

First, that no one country's assessment system was perfect either in design or in outcome. Second, that good assessment doesn't happen by chance; it has to be carefully planned for and developed using the key conditions of content standards, cognitive operations, cognitive demand and question type.

A third factor that could arguably be added and which has been the subject of recent controversy is the rapidity of change. As the regulator has argued, when a qualification and thus its assessment is frequently changed, the risks to security of standards and comparability of standards over time become greater.

Next steps

These and other challenges are currently under consideration as part of our challenge to develop world class qualifications. For more information about the research into World Class Qualifications and further developments, please visit [World Class Qualifications](#) or contact wcq@pearson.com.