

Pearson World Class Qualification

The Evidence Base for Tiering in GCSE Exams

Contents

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

The issue

Essentially, it's a system whereby different levels of exam are offered within a single overall template. In most cases, two tiers have been available: higher, covering GCSE grades A*–C, and foundation, covering grades C–G. The thinking behind tiering is to ensure that a particular exam can be offered to as wide a range of student abilities as possible, thereby ensuring some reward for efforts.

The problem, however, is that the application of tiering means judgements have to be made about a particular student's potential before they are entered for an exam and such judgements may be neither fair nor accurate. In addition, by entering a student for the lower tier, their achievements are restricted to that particular grade boundary.

Tiering therefore raises important issues about exam design and challenge on the one hand, and teacher judgement and equity on the other. Michael Gove has expressed his opposition to tiering and has called for the practice to be abandoned in his latest proposals for the reform of GCSEs.

What does the research tell us?

As exam design has developed over the last 30 years or so in pursuit of differentiation, so a body of research has emerged which offers some useful insight. Overall, four broad themes emerge from this research:

1. A **range of different approaches** has been proposed over the years to try and allow for differentiation. While tiering has become the standard model, others have attracted attention. These include:
 - **Ranking:** using raw tests from two different ability groups, each of whom had sat two papers, one common and one different, and combining the scores to create an overall ranking.
 - **Scaling:** a similar approach to ranking but using an equal mean and standard deviation for each ability group to create a scaled score.
 - **Regression:** again using two groups of scores but using these to estimate the score they would have achieved on a third paper and then collating this to derive a ranking.
2. Tiering is **more popular** in some subjects than others. It has, for example, been widely used for mathematics but not for history, where differentiation has been achieved on the basis of outcome. Under this, students of all abilities sit a common paper but are graded on the basis of differentiated responses, something that is inappropriate for a subject such as Mathematics. Different attitudes to tiering were further evident in the Government's 2012 consultation on reform of Key Stage 4 qualifications. Responses from subject associations showed strong support for tiering in subjects such as Maths and Science, variable support from Modern Languages and Geography and opposition from History experts.
3. **International evidence** is useful but not readily transferable. A number of countries employ a system of differentiated exams with similar characteristics to GCSEs. These include Hong Kong, Korea and Finland, each of which has achieved eminent results on PISA tests, thereby suggesting that the practice of differentiated exams could have a positive effect. However, not all countries undertake formal exams at age 16 and only a few of these adopt such a high-stakes approach to these exams as in this country. It is therefore difficult to draw conclusive evidence from international evidence.

4. The **social and motivational impacts** of tiering on students are important to consider but can be hard to disaggregate. A lot of research has been conducted in this area, looking at such issues as the impact on girls as compared to boys, teacher perceptions and value judgements, and the role of socio-economic status generally. There is some evidence, for example of students entered for a lower tier losing motivation, equally of teachers being swayed by stereotyping and a desire to play safe, but these can be hard to prove conclusively. What does seem clear is that such issues would remain however exams were designed and tend to be exacerbated by the continuing emphasis on high-stakes testing.

In summary

Clearly, however the exam system is designed the tension between being able to measure and record achievement fairly, reliably and validly on the one hand while providing encouragement, stimulation and access on the other, remains delicate.

The research that Pearson has conducted as part of its work in developing World Class Qualifications has revealed a complex picture, so the challenge is to develop assessment approaches that conform to political wishes and regulatory guidelines whilst meeting subject needs and providing suitable challenges for different ability levels within secure assessment structures.

Next steps

For more information about the research into World Class Qualifications and further developments, please visit **World Class Qualifications** or contact wqc@pearson.com.

