



The Mathematical Association  
Supporting Mathematics in Education

On the subject of **Measuring Performance** we say that there should be a change of emphasis ...

**The Mathematical Association set out its approach to measuring performance, and specifically to the use of league tables, in a position paper on assessment in 2002. A decade later, we are pleased to see the British Academy publishing near-identical recommendations. We believe that there should be a focus on reducing opportunities for the 'gaming' which so readily distorts students' experience of learning mathematics, on alleviating stress in students and teachers alike, on improving performance within schools rather than making comparisons between them, and on taking better account of prior learning.**

## 1. A long-held position

The Mathematical Association has long argued for the insightful and sensitive use of performance data and league tables. Our 'Position Paper on Assessment', written by Doug French in 2002, acknowledged that 'society does have to measure the performance of the education system and provide certification to indicate the attainment of individual students, but this should be done in ways which ensure that the long term goals of education are not compromised.' The Association's position was outlined in the points given in the appendix to this statement. In March 2012, the highly-respected British Academy published a report on *Measuring Success: League Tables in the Public Sector*, in which its recommendations for education, also given in the appendix, accord with our own to a high degree.

In the decade between the publication of these two documents, The Mathematical Association provided further statements. We joined with ATM and NANAMIC in 2003 to produce a report on 'The Impact of Assessment on Learning and Teaching Mathematics', a report which highlighted how performance measurement drives many schools to focus on the C/D borderline at the expense of both stronger and weaker students. We commented that the concentration of effort and resources on these students:

were considered to be both inappropriate in distorting the learning of individual students and unjust in leading to the relative neglect of other students. Whilst it is desirable to help students to achieve at the highest level possible, a sensible balance needs to be struck between providing effective learning, which is useful in the long term to students, and achieving a particular grade in order to satisfy the targets set for schools, departments and individual teachers. It is clear that many teachers feel that pressures are making such a balance impossible to achieve.

Then in 2006, in a 'Response to the House of Lords Select Committee on Science and Technology - Mathematics Teaching in Schools', The Mathematical Association argued that:

The current system of assessment and all the accompanying targets and league tables are having substantial ill effects on the teaching and learning of mathematics. A radical shift away from the current dominance of tests, examinations, targets and league tables is essential if standards in mathematics are to be improved.

The excessive emphasis on tests and examinations has the very serious effect of skewing all classroom activity towards the short term goal of maximising test results. Important aspects of mathematical learning that are hard to assess become optional in the eyes of both students and teachers. A substantial reduction in statutory testing is necessary if the quality of teaching and learning is to be improved and would contribute greatly to creating an environment in which mathematics teaching is a more congenial task.

The comments we have made over the decade appear to be the very conclusions that are being made in the latest reports and documents coming from various bodies. So we now contend that there needs to be a belated shift of emphasis in the regimen of performance measurement and we lay out below where the defects in the current approach lie and how a more equitable and pedagogically-sound system might be put in place.

## **2. Performance rankings, their purposes and weaknesses**

There is a range of statistics gathered and published on the performance of schools. They include the percentage of students attaining 5 A\*-C grades including mathematics and English, 5 A\*-C grades in the English Baccalaureate and the three levels of progress between KS2 and KS4. The contents of the basket of statistics collected are under constant review, with 'statements of intent' to adjust the régime published by the DfE as required.

According to Foley and Goldstein<sup>1</sup>, the purposes of performance rankings, such as school league tables, are threefold, to:

- provide public accountability, 'whereby those who provide resources to run ... schools can form judgments about where improvement is needed or particular action is required',
- 'enhance performance by encouraging competition between institutions in a quasi-market environment', and thereby inform parental choice of schools,
- afford a measure of control by 'providing targets, such as those associated with school examination results', which is 'seen as a powerful means of making policy indirectly by providing appropriate incentives for behavioural change.'

The same authors cite four deficiencies or adverse effects, which for education, are:

- behavioural change that can be characterised as 'gaming' — schools seeking to 'improve their ranking, by manipulating exam entry policy to the detriment of student choice, or even by excluding low achievers',
- that 'the range of what is measurable and hence amenable for use in performance indicators is limited, and concentration on these detracts from other, less quantifiable objectives such as breadth of learning',
- that league tables need to be contextualised, specifically, 'school examination results need to be adjusted for the intake achievements of students when they start at a school – so called 'value added' rankings',
- that 'the uncertainty surrounding any given ranking is very large, and in many important cases so large that no statistically meaningful comparisons can be made, nor can useful user choices be sustained.'

## **3. Specific Concerns**

---

<sup>1</sup> Beth Foley & Harvey Goldstein, *Measuring Success: League Tables in the Public Sector*, British Academy Policy Centre, March 2012.

The Mathematical Association's mission is to promote the highest quality of learning and teaching in mathematics and consequently we have a number of specific concerns.

### 3.1 'Gaming'

For some time, one of The Mathematical Association's main concerns has been that league tables skew what is happening in classrooms. This skewing takes many forms including a loss of continuity and progression in mathematics and the repetitive practising of a limited number of skills at the expense of encountering deeper concepts and the development of the full range of mathematical ways of engaging with the world: the range of valued knowledge, skills, concepts and behaviours, outlined by ACME in its report, *Mathematical Needs*.<sup>2</sup> But it is somewhat inevitable wherever metrics and targets exist, simply because heads of department, head teachers and parents expect that students will perform well and be seen to perform well. This introduces the prospect of 'gaming', whereby the teacher's focus is on those aspects of learning that are measured whilst other aspects are neglected. In her *Review of Vocational Education ('Wolf Report')*, Alison Wolf argued that the current performance table system has 'created perverse incentives', and we note the strength of the language used.<sup>3</sup> So, it is the high-stakes nature of universal external testing that is the damaging factor here.

### 3.2 Curriculum distortion

Where is the gaming which so distorts students' experiences of mathematics learning most keenly felt? It occurs at many stages, including at primary school, where teachers are expected to show improvement in performance for each child that they teach. This often leads to various strategies to enhance attainment measured at the end of the term or year. The next teacher finds that many children are not really confident at the level they are said to have attained, but improvement is expected to be shown. This goes on year after year. The teacher who sees the need to fill in gaps at lower levels or to consolidate and develop understanding before moving on may well find that their pupils appear to have made less progress than expected, even though they may well be better prepared for the mathematics they will meet in the future.

In the upper years of primary school, the prominence of league tables leads to the skewing of much of the teaching, so that mathematics (with English and, to some extent, science) tend to dominate what is delivered through the session until the Key Stage 2 tests. Thereafter, these subjects, which all benefit from regular coverage, often get very little teaching time, whilst the focus shifts to the subjects that have been neglected. This hinders the smooth and confident transition to secondary school, as very little mathematics may have been studied for several months. As The Mathematical Association argued recently in our *Evidence on the National Curriculum Review* (2012, §31):

National Tests at the end of Key Stage 2 hinder the smooth transition to Key Stage 3 as many learners spend much of Year 6 being coached for these examinations instead of deepening their mathematical skills further. The emphasis should always be on teaching children mathematics rather than teaching them to pass mathematics examinations.

In secondary schools in recent years, mathematics has been taught in such a way that emphasis is placed on acceleration to higher National Curriculum levels and examination grades rather than on depth of understanding, i.e. on narrow and short-term 'teaching to the test' rather than teaching for understanding and as a foundation for future learning. Those aspects of mathematics which are easy to test have been tested and as a consequence they

---

<sup>2</sup> ACME, *Mathematical Needs: The Mathematical Needs of Learners*, June 2011, Section 3.

<sup>3</sup> Alison Wolf, *Review of Vocational Education — The Wolf Report*, DfE, March 2011, p.136.

have been given undue prominence at the expense of other topics and opportunities to use and apply mathematics.

Perhaps, gaming has been at its most evident in the presentation strategies adopted as GCSE approaches. In order to meet targets, there has been widespread use of early and repeated presentation for GCSE Mathematics, with students being entered according to a variety of strategies, which have resulted in a repeated cycle of failure and revision for some, and success and neglect for others as their focus and their teacher's focus move elsewhere. There has been a climate in which the popularity of modular examinations has become inevitable, with 'teaching to the test' for each module, rather than teaching for understanding, connecting the learning and laying down foundations for later learning. With so much gaming taking place because of the prominence of performance measurement, Ofsted's recent conclusion comes as little surprise and will be welcomed by many:

While the best teaching developed pupils' conceptual understanding alongside their fluent recall of knowledge, and confidence in problem solving, too much teaching concentrated on the acquisition of disparate skills that enabled pupils to pass tests and examinations but did not equip them for the next stage of education, work and life. ... Monitoring of each pupil's understanding was not strong enough to ensure that pupils learnt and progressed as well as they could.<sup>4</sup>

### **3.3 Threshold attainment, understanding and future progress**

The Mathematical Association wishes to emphasise that success on an assessment does not necessarily equate to full understanding, nor is it a guarantor of future progress. We are fully supportive of Ofsted whose recent report, *Mathematics: Made to Measure*, repeatedly makes the point. Inspectors found that results often flatter candidates' understanding, commenting that:

Quick-fix approaches were particularly popular. Aggressive intervention programmes, regular practice of examination-style questions and extra provision, such as revision sessions and subscription to revision websites, allowed pupils to perform better in examinations than their progress in lessons alone might suggest.<sup>5</sup>

They also noted that despite the gaming strategies adopted,

the percentage of pupils not reaching the expected level or grade for their age increases as pupils progress through their mathematical education, and is more marked for some groups than others. This suggests, strongly, that attaining a key threshold does not represent adequate mastery of skills and sufficient depth of conceptual understanding to prepare pupils for the next stage of mathematics education.<sup>6</sup>

And specifically with regard to measures designed to increase the number of A\* to C passes, these tactics ... account for the rise in attainment at GCSE; this is not matched by better teaching, learning and progress in lessons, or by pupils' deeper understanding of mathematics.<sup>7</sup>

### **3.4 Impoverishing the curriculum of the most able and the least able**

---

<sup>4</sup> *Mathematics: Made to Measure*, Ofsted, March 2012, tenth key finding, p.9.

<sup>5</sup> *Ibid.*, §39, p.19.

<sup>6</sup> *Ibid.*, seventh key finding, pp.8-9.

<sup>7</sup> *Ibid.*, §40, p.19.

An issue that arises specifically in regard to schools' preparation for GCSE is that every effort is made to maximise the proportion achieving at worst a grade C. The ability to just reach this threshold is characteristic of moderately-talented students. It is they for whom special support is very often provided, whilst those of higher and lower ability are neglected. To give just one example, students have sometimes been presented early for mathematics and/or English GCSE, with failure to achieve a grade C in one of them resulting in all the remaining mathematics and English lesson-time being assigned to the that subject. Here we have a lack of continuity and progression for able students. As Alison Wolf commented:

The '5 GCSEs' measure incentivises schools to neglect those at the top of the attainment range, because like any measure with a threshold it focuses attention on students just on one or other side of it. It is important that schools be given a strong incentive to pay attention to their least academically successful *and* their most academically successful pupils.<sup>8</sup>

With his eye firmly on the most able, Sir Michael Wilshaw commented recently that 'too many of our able pupils do not fulfil their potential'. He added:

The extensive use of early GCSE entry puts too much emphasis on attaining a grade C at the expense of adequate understanding and mastery of mathematics necessary to succeed at A level and beyond. ... Our failure to stretch some of our most able pupils threatens the future supply of well-qualified mathematicians, scientists and engineers.<sup>9</sup>

And although here we have been discussing only one performance measure, the argument is equally valid for any high-stakes measure: gaming is inappropriate on grounds of equal opportunity and damaging except sometimes for those in the focus group.

### **3.5 Appropriate intervention: support for all**

Students underperform for a variety of reasons, including home circumstances in which learning is not valued, lack of application, meagre conceptual understanding, poor attendance and ineffectual teaching. Whilst it is important to target groups for intervention, this should not be restricted to those whose likely improvements will have the greatest effect on the department's or school's performance data. We are pleased to see Ofsted providing an example of best practice in its report, *Mathematics: Made to Measure*, §220, introduced in these terms:

The scope of intervention in the secondary schools varied, with some intervening with all those at risk of underachievement and others still focusing mainly on pupils at risk of narrowly missing the Key Stage 4 threshold target of five or more GCSEs at A\* to C including English and mathematics. The most equitable practice focuses on all pupils who are underachieving.<sup>10</sup>

Ofsted defined 'prime practice' as being the 'intervention for all who need it' they found in one school:

Intervention and revision contributed significantly to pupils' success in examinations. Pupils were divided into key groups: low to middle ability girls who had underachieved previously; underachieving boys; poor attenders; those on track to meet the five A\* to C threshold; and a group who were making secure progress whatever their starting points. All groups

---

<sup>8</sup> Alison Wolf, *Review of Vocational Education — The Wolf Report*, DfE, March 2011, p.136.

<sup>9</sup> Sir Michael Wilshaw, *Mathematics: Made to Measure*, Ofsted, March 2012, p.4.

<sup>10</sup> *Ibid.*, §220, p.84.

received support and encouragement relevant to their needs. In this school, intervention was about the achievement of individuals rather than simply those on the C/D borderline.<sup>11</sup>

The Mathematical Association believes that this is the only equitable approach.

### **3.6 Consequences for teachers, students and parents**

The pressure on teachers to attain expected targets creates unnecessary stress in teachers. This may be related to anxieties about job security. They may also feel frustrated that their teaching is adversely influenced by managers, bureaucrats, and politicians whose primary emphasis may not always be the quality of students' learning. The effect of pressure on students to attain the expected levels is often a sapping of self-confidence, even where that pressure is self-imposed. The situation is made worse where the anxiety felt by teachers is passed on, even unwittingly, to students. And yet, inspectors have found that:

leaders in the primary and the secondary schools were driven strongly to improve results in mathematics, clearly influenced by the public profile of their results and how those results are used. Since the last survey, it has become increasingly common for teachers' performance management targets to include measures of attainment and/or progress for the pupils that they teach. Thus, accountability has increased.<sup>12</sup>

Teachers who are prepared to re-teach skills reportedly already acquired may well be seen as unsatisfactory and be creating problems for themselves. It is in the interest of teachers to ensure that in the short term attainment shows significant progress, but this is often not in the interests of the children. The nature of mathematics is such that it is often necessary to re-visit or back track and in some cases go back to the beginning and start again. Unless teachers are encouraged to do this and can be confident that apparent lack of progress will not be seen as a cause of concern by school leaders, parents and inspectors, we believe that significant improvements to real attainment will not be made.

Parents may be misled by superficial reading of performance tables. They may be unfairly critical of a school and so create unhelpful relationships. A parent's loss of confidence in a school does not help a child. Performance tables do not provide information about important qualities of a school, e.g. ethos, which is available only through Ofsted reports.

## **4. Seeking a best approach to measuring performance**

The Mathematical Association would argue that the measurement of performance (whether expressed in such a way or through terms such as 'appraisal' or 'objectives for improving provision') brings certain pressures to bear on schools and on teachers, some no doubt positive and others which have clear, adverse effects on teaching and learning mathematics. We recommend that there should now be a change of emphasis.

1. Performance measurement should be used to support:
  - students to understand the quality of the range of their mathematical skills and, with the support of their teachers, identify their next areas for development,
  - schools to celebrate success and to identify and take measures to counter underachievement.
2. Performance tables should be based upon a rich mix of metrics that between them cover, as fully as is possible, what we mean by a good (mathematical) education. This

---

<sup>11</sup> Ibid.

<sup>12</sup> Ibid., §121, p.51.

would ensure that the tables could not be gamed; the only way to climb up the tables would be to improve the education provided.

3. Consideration should be given to withholding performance tables from the public, or else limiting their extent and educating the public and the media as to how to guard against their misuse and misunderstanding.
4. National performance should be monitored by an independent body using sampling techniques, backed by evidence gathered from the classroom by Ofsted.

### **Appendix: Recommendations of The Mathematical Association (2002) and the British Academy (2012)**

#### **The Mathematical Association's 'Position Paper on Assessment', Doug French, 2002**

1. Monitoring national and school performance requires a very much lighter touch and should be done much less publicly;
2. National performance would be monitored much more effectively, and acceptably, if done by an independent body using sampling techniques, as used to be the case with the Assessment of Performance Unit (APU), together with using more qualitative evidence obtained through a sensitive system of inspection whose priority was to offer immediate and specific advice for improvement;
3. Targets for schools, LEAs and individual teachers based on proportions of students achieving particular levels or grades, and the related league tables, should be abandoned: schools and teachers need constant encouraging advice on how to improve the quality of students' learning rather than exhortation and pressure to maximise test results;
4. The system requires a substantial change of emphasis from summative to formative assessment and from a focus on easily testable skills to embracing much wider and more long term goals.

#### **Beth Foley & Harvey Goldstein, *Measuring Success: League Tables in the Public Sector*, British Academy Policy Centre, March 2012.**

1. The linking of league tables to rewards should be weakened to reduce the side effects of inappropriate 'gaming' and to reduce stress among teachers, parents and students. This would also have the desirable effect of making the results a more objective evaluation of performance. The problematic consequences for schools serving the most disadvantaged pupils particularly need to be addressed.
2. The government should consider ways to prevent league tables being exploited by the media, such as ensuring that measures of uncertainty are provided around any institutional results. Associated with this there could be a campaign to better inform the public at large about the strengths and limitations of league tables, although any such attempt poses considerable challenges.
3. Consideration should be given to alternative ways of using quantitative information to monitor educational performance generally. This can be achieved by in-depth study of a sample of schools and students within a national database. A useful model is the Assessment of Performance Unit that was set up in the 1970s in England and discontinued in the 1980s.
4. Consideration should be given to using performance information as a screening device rather than publishing as league tables, as in the Hampshire experiment. This could be accompanied by an emphasis on evaluation and inspection systems that are designed to emphasise ways of assisting schools to cope with problems rather than 'exposing' them using public rankings.
5. Ways to rely less on a small number of indicators should be sought, as well as those which cover more aspects of learning.
6. More appropriate statistical analysis models should be used to describe institutional differences that allow for differential performance for different groups of students. In particular, there should be a shift away from the comparison of individual institutions towards research that helps to identify modifiable factors that appear to be related to good performance.

7. An ethical code to govern the publication of school performance measures should be formulated ... This would be based on two broad principles: that unjustified harm to those to whom the information applies should be prevented, and that there should be no *absolute* publication rights for performance data.
8. Further consideration needs to be given to the role of inspection and accreditation agencies as a means of evaluating individual institutions. Trust in such agencies may not be easy to achieve, especially when they are perceived to be instruments of government.