HMI Ofsted report for Mathematics 2008 or why teenagers are maths dunces[1]

Stephanie Prestage and Pat Perks

School of Education, University of Birmingham

This paper offers a brief analysis of the HMI mathematics report published on the 19th September 2008 (Ofsted, 2008). We consider alternative discourses for the data used by the authors and for their explicit and implicit messages with the resultant blame laid on teachers. The serious message of this paper is the startling nature of the attribution of blame after more than a decade of prescription and national targets which are regularly inspected.

Key words: Inspection, blame, performance culture

Background

Much has been written about the compliance culture created by regulation and audit. Management teams at all levels are required to set clear directions and are to measure the value and effectiveness of practices especially those of teachers. But as Avis (2003) points out:

> Performativity through its chain of targets and accountability operates with a 'blame culture' where accountability becomes a means by which the institution can call to account its members. … In this way performance management operates as a mechanism to discipline members so that they follow preferred directions and practices. (324-325)

Many authors claim further, that centrally determined targets of anything but the broadest are unlikely to raise standards and in fact may not help to sustain improvement (Hargreaves, 2003; Sachs, 2003; Stronach et al 2002). Moreover these authors also believe that use of such targets is most likely to encourage conservative practices and suppress innovation.

It is interesting to note that the most influential mathematics report in the early 1980s (Cockcroft, 1982) commenting on whether there was a need for teaching methods to be detailed decided that: "we are aware that there are some teachers who wish us to indicate a definitive style for the teaching of mathematics, but we do not believe that is either desirable or possible." Ten years ago, Brown (1999) highlighted the growing pedagogic control, suggesting that:

> The post-war progressive era of teacher and pupil autonomy is apparently over, and the education system is to be driven by national targets and norms and regularly inspected like steel production in the soviet state or rice production under Chairman Mao. (15)

The National Strategy provides a great deal of advice in the form of 'distance education' materials –folders, CDs, videos, training packs etc. Teachers working across the educational spectrum are expected to interpret these suggestions, to make sense of them within their classrooms. The Strategy endorses a high quality

[1] Daily Mail, 19th September 2008

formative assessment interacting with pupils to make clear what the objectives are assessing how far they have progressed in achieving them and using information. This is seen by Brown (1999) as:

> …the tightest ever control by government on mathematics, with central prescription not only of national curriculum and national tests, but also of teaching styles. (374)

So after a decade of increasing specification, audit and inspection, what is happening in mathematics education? The results from Ofsted reveal disappointment but where is blame apportioned?

Method

The data source for this paper is the HMI report itself. We adapt a framework offered by Jupp and Norris (1993) to analyse the document. This framework is given as a set of four questions shown in the frame below

1. Who writes these discourses and whom do they represent or purport to represent?
2. What is the intended audience of such writing? 3. What does critical reading of these documents tell us about what is seen as problematic; what explanation is offered for what is seen as problematic? 4. What alternative discourses exist?
adapted from 50

Q1. Who writes these discourses and whom do they represent or purport to represent:

Clearly the authors of the report are the mathematics HMI OfSTED inspectors, representing an independent government quango. Our emphasis here would be on the word "independent." However as the discussion below reveals it is only teachers who are blamed for the observed underperformance of students. All other agencies are enjoined to help to 'fix' the teachers, which seems to offer a particular political agenda with an undermining of independence.

Q2. What is the intended audience of such writing?

This is ambiguous, in that such reports are usually for the education community at large. But we suggest that the main audience is the DCSF since the advice to teachers of which there is plenty comes with explicit admonishment. Is this to justify an increasingly expensive inspection machinery and to offer newspaper headlines to allow a Government to step in with more quick fixes to make more headlines?

Q3. What does critical reading of these documents tell us about what is seen as problematic; what explanation is offered for what is seen as problematic

The report reveals much that is problematic and highlights the consequence of emphasis on increased test results at all ages creating a culture of 'teaching to the test', with its allied fault finding of teachers and teaching. Yet the nature of the evidence is never seen as problematic. Perhaps Ofsted inspectors have to be oversure of their interpretations in order to do the job, but the report ought to have been more intellectually honest and rigorous.

Q4. What alternative discourses exist?

The anecdotes which are offered as evidence are based on the observations of experienced observers but the discourse in the report allows for only one

interpretation and one scapegoat (the teachers) and yet the evidence can and needs to be challenged by alternative interpretations of the findings.

Discussion

Several themes emerge from an analysis of the text but we choose three in particular to write about here.

Those important test results

The report is organised into four sections, the executive summary, key findings, recommendations and then the main text. It is most likely that, given the length of the report that most readers will read the first three and perhaps parts of the main section. It would seem important, therefore, that the first two sections are an accurate reflection on the data. So, what is the first key finding? Surely what is placed first is the most important, the most significant (lists by default create a hierarchy). The first *key finding* reminds all readers of the importance of those improved test results:

> Results of national tests at Key Stages 2 and 3 and GCSE examinations have shown an upward trend for several years (6, Key findings)

As this is the first key finding one might presume that the teachers are about to be congratulated for their hard work and effort. No such thing. Tests results are better than ever before but teachers are still considered to be at fault. They are criticised for getting better results because this has affected their pedagogy. The executive summary reveals some problems and the inherent problem of these improved and increasing results is to be 'blamed' on teaching to the test.

> Evidence suggests that strategies to improve test and examination performance, including 'booster' lessons, revision classes and extensive intervention, coupled with a heavy emphasis on 'teaching to the test', succeed in preparing pupils to gain the qualifications but are not equipping them well enough mathematically for their futures. (4, Executive Summary)

Are the inspectors concerned about this? Well, yes. There is criticism of "teaching to the test", alongside the suggestion that booster classes, revision classes and excessive intervention so that they may pass examinations, are leading to a narrow and joyless experience for many pupils. Yet there is nothing in the key findings. You have to go to page 44 within the report to find a rather critical statement about the consequences of teaching to the test:

> … many concentrated on approaches they believed prepared pupils for tests and examinations, in effect, 'teaching to the test'. This practice is widespread and is a significant barrier to improvement. (44, §111)

This 'significant barrier to improvement' gets no mention in the key findings nor in the recommendations. Barbour (2008) also noted with concern that the inspectors have nothing to saying about this problem in the key findings

> … expecting to see some robust advice to government to free teaches from the treadmill of high stakes testing. And guess what… there is no such recommendation. (Barbour, 2008)

In relation to testing and attainment, the inspectors also have something to say but oddly, for mathematicians, they use numbers carelessly, assuming learning and

progression to be linear and that level attainment at KS3 is meaningful, valid and predictive of performance at KS4:

> Key Stage 3 test results are improving and a greater percentage of pupils reach the vital threshold of grade C at GCSE level, but this does not tell the whole story. Based on the gains made at Key Stage 3, more pupils than at present should be reaching the higher GCSE grades (4, Executive Summary)
>
> The improvements made in Key Stage 3, however, are not built on sufficiently during Key Stage 4. Indeed, pupils' progress during Key Stage 4 has declined over the past few years. In 2007, 79% of pupils who had reached Level 6 at Key Stage 3 went on to pass GCSE at grade C or higher, and 26% did so from Level 5. (10, §4)

Wiliam (2001) explains that even with a reliability score of 0.8 on the outcomes of the KS3 tests, which is quite normal, 43% of students could be misclassified. The percentages quoted become easily deconstructed into meaningless and uninformative numbers.

Validity and reliability of the data

The report purports to be evidence based yet the reliance on personal interpretation by an observer, however skilled, which fails to take account of context and influences on behaviour brings such data into disrepute. Reading the main findings of the report you will find many places where inspectors are negative in their reports on teachers. The use of language is pejorative and assumes the observations of lessons to represent a truth that can be generalised.

> A substantial amount of teaching is no stronger than satisfactory and, in these lessons, pupils do not learn as quickly as they might. (19, §26)
>
> A different starting point might have been to use an easy example ….. Learning would have been better if ….. (20)

Assumptions abound. Have you ever talked to teachers about their anxiety when being observed by an Ofsted inspector? The lesson being observed is not 'normal', there is tension, nervousness, angst.

> When the inspector asked her (a pupil working with $A=\pi r^2$) …a question… she immediately assumed she was wrong, as she was not used to being asked to interpret her answers. .... The pupil's understanding would have been better if the teacher had …. (54)

The certainty of the discourse does not permit alternative interpretations, such as the pupil being unwilling to answer in a lesson observed by an outsider, that her understanding at that moment would not be enhanced by such a process, or that asking for interpretation was not a normal activity for that teacher. The evidence is not seen a problematic and open to alternative discourses. Even worse, generalisations to wider teaching situations are made from a single quoted observation:

> **Commonly**, teachers remain at the front of the classroom during starter activities, while introducing a topic and during class discussions. This means they miss important information about the questions pupils find difficult or too easy and do not recognise where an early slip is interfering with pupils' learning. For example, many pupils draw axes with unequal spaces between units which prevent them from plotting straight line graphs correctly, yet they continue to work on them unnoticed for too long. (120, c.f. example of a lesson on page 19)

As teacher educators we are familiar with our student teachers not directing movement or looks to the part of the room where we are sitting. An alternative discourse on this anecdote allows for observation by an outsider to change behaviour (in an attempt to divert the stress of being observed) and that one observation does not indicated usual behaviour and certainly cannot be evidence for this to be considered as event occurring "commonly".

Schemes of work

There are statements in the report which can be interpreted as contradictory. For example, two key findings are given in relation to schemes of work. The first is that schemes of work are generally poor.

> Schemes of work in secondary schools were frequently poor, and were inadequate to support recently qualified and non-specialist teachers. (6, Key Findings)

However following on from this statement is the comment that management and leadership was satisfactory and getting better

> The quality of leadership and management of mathematics was good or better in 71% of the primary schools and 51% of the secondary schools visited, although it has improved in the latter in the last two years. (6, Key Findings)

How can this be? Schemes are work are frequently poor and yet leadership and management are good or better in over half of the schools. If management and leadership are responsible for working with their teaching teams on schemes of work, how can they be judged as good or better when they produce poor schemes? In the depths of the report more information is found about the key finding, being particularly damning for secondary schools.

> Good schemes of work were rare in secondary schools. It was not uncommon for teachers to use only examination specifications and textbooks to guide their lesson planning, focusing on content rather than pedagogy. Few schemes included guidance on matters such as the most effective teaching approaches, how to meet the full range of pupils' needs or on what constitutes an appropriate level of challenge. They provided insufficient support for teachers who were at an early stage in their professional development or for staff who were not mathematics specialists. (25, §46)

Now our experience of schemes of work in secondary schools (40 PGCE secondary students working in secondary schools each year) is their use of the medium term plans from the Strategy team. Indeed the report finds the same.

> In many secondary schools, apart from adaptations needed because of changes in examination specifications, there has been little progress in developing the mathematics curriculum since the Key Stage 3 Strategy's sample medium-term plans several years ago. (25, §47)

So, and here is our confusion, schools are using the medium terms plans from the Strategy, many of these schools have been inspected and schemes of work classed as poor. Worse still, given that the inspectors find those based on medium terms plans poor, who is blamed for their use? Why, the teachers of course. And whose advice is sought? Who else but the Strategy! The Recommendations ask for the Strategy to:

> … provide guidance for schools on enhancing subject expertise in mathematics …. [and] devise guidance for teachers on the effective use of mathematics-specific pedagogy to aid the development of pupils' understanding (7, Recommendations)

If the medium term plans are not considered sufficient for good schemes of work and advice from documents and consultants have not prevented 'teaching to the test', and have encouraged booster classes etc., is the really Strategy qualified to improve the situation?

Conclusion

The charges which the chief HMI levels against teachers are serious and wide ranging - poor subject knowledge, poor pedagogic content knowledge, poor management and poor schemes of work. If she had named a particular school or particular teachers she might have had a series of libel actions against her name. Instead she generalised from the individual observations and as a result all mathematics teaching was covered in the mud of a particular interpretation of the 'evidence'. The press had a field day.

> "Ofsted criticises maths lessons" Guardian (19/09/2008)

> "Teaching style turns children off maths say inspectors" Times (19/09/2008)

> "Schools failing pupils on maths" Daily Express (19/09/2008)

> "Half of maths lessons not good enough" The Independent (19/09/2008)

> "almost half of the lessons are no better than satisfactory" The Telegraph (19/09/2008)

> "Why teenagers are maths dunces" Daily Mail (19/09/2008)

Mud smeared anonymously is however still mud smeared – and all the more difficult to wash away. Quite what constructive purpose could possibly be served by such irresponsible and wholesale scapegoating of teachers is beyond our comprehension. We suggest those in charge look upward and within. Their performance is poor, their political acumen naïve and their use of data misleading.

References

Avis, J., 2003. Re-thinking trust in a performance culture: the case of education, *Journal of Education Policy* 18[3]: 315-332

Barbour, R., 2008. *One Cheer for Ofsted*, The Mathematical Association Newsletter, October 2008, 137

Brown, M., 1999. Swings of the pendulum. In *Issues in Teaching Numeracy in Primary schools,* ed. I. Thompson, Milton Keynes: Open University Press.

Cockcroft, WH., 1982. *Mathematics Counts*. London: HMSO

Hargreaves, A., 2003. *Teaching in a knowledge society: education in the age of insecurity,* Milton Keynes: Open University Press.

Jupp, V. and C. Norris. 1993. Traditions in documentary analysis. In *Social Research: Philosophy, Politics and Practice,* ed. M. Hammersley, London: Sage.

Ofsted 2008. *Mathematics: understanding the score*, London: Ref. No. 070063

Sachs, J., 2003. *The activist teaching profession* Milton Keynes: Open University Press.

Stronach, I., B. Corbin, O. McNamara, S. Stark, T. Warne. 2002. Towards an uncertain politics of professionalism: teacher and nurse identities in flux, *Journal of Educational Policy* 17[1]: 109-138

Wiliam, D., 2001. Reliability, validity and all the jazz, *Education 3 - 13* 29[3]: 7-21