FORUM
Volume 59, Number 2, 2017
www.wwwords.co.uk/FORUM
http://dx.doi.org/10.15730/forum.2017.59.2.121

EDITORIAL

Assessment: crisis and resistance

TERRY WRIGLEY

May 2016 was the first time that primary school SATs tests were based on Michael Gove's revised National Curriculum, introduced in 2014. The new curriculum set targets without any thought to children's development; it was simply a race to become a global winner at PISA. A hundred academics, including some associated with this journal, warned publicly in 2013 that the new curriculum would be disastrous. Their letter hit newspaper front pages under the headline 'Too Much Too Soon'. Gove paid no heed to the warning, preferring to denounce the academic experts as 'bad academics' and 'enemies of promise'.

As the 100 academics' open letter correctly predicted, it put pressure on teachers to rely on rote learning and endless test preparation, caused untold stress to children, and led to a sense of failure and demoralisation. The letter explained why it would be counterproductive in its declared aim of raising standards. Indeed, 2016 saw almost half of 11-year-olds transfer to secondary school bearing a failure label in at least one of reading, writing and mathematics. A third were failed in reading alone. The situation was even worse for children in poverty: of those currently entitled to free school meals, two thirds were failed in at least one of these tests.

Parents became angry and upset at what was happening to their children, and, following a call across social media, a hundred local groups took their children out of school for a day just before the 2016 tests. The 'More Than a Score' coalition was formed in autumn 2016, and included the National Union of Teachers (NUT), parents, major curriculum bodies such as the National Approach for Primary Education (NAPE) and the United Kingdom Literacy Association (UKLA), and the British Educational Research Association (BERA), representing educational researchers. Unfortunately, there was not time to build sufficient momentum for a nationwide boycott in 2017.

Under the weight of public concern, the House of Commons Select Committee initiated an investigation, receiving hundreds of critical responses. Its final report (March-April 2017) was damning. It concluded that the Standards and Testing Agency did not oversee implementation of the new system effectively, causing major disruption. The tests were poorly designed. The writing assessment focused overwhelmingly on technical aspects of spelling, punctuation and grammar, neglecting creativity and composition.

The MPs (both government and opposition) called for the stakes to be lowered:

> Assessment is closely linked to the accountability system in primary schools, with Key Stage 2 results used to hold schools and teachers to account on the progress and attainment of pupils. However, the high stakes system can negatively impact teaching and learning, leading to narrowing of the curriculum and 'teaching to the test', as well as affecting teacher and pupil wellbeing.

It condemned the sudden increase in difficulty in no uncertain terms:

> We agree with the Government's aim of raising standards at primary school but think that setting extremely challenging targets only leaves many students feeling they have failed, when in a previous year they would have succeeded. Expected standards should be raised over a much longer time period to give schools a chance to adjust to new expectations.

It also raised serious question marks about the government's desire to reintroduce baseline testing at the start of reception year.

There is much to build on here. *More Than a Score* had launched with a national conference in London in December 2016. A regional conference organised by NUT branches, parents' groups and *Reclaiming Schools* (a network of researchers formed to support various NUT campaigns) took place in Manchester in October 2016, followed by conferences in Newcastle and York under the *More Than a Score* umbrella. The movement's resources include two pieces of research commissioned by the NUT – 'Exam Factories' and 'They Are Children... Not Robots' – as well as the *Reclaiming Schools* collection 'The Mismeasurement of Learning'. Finally at least one head teacher, Jill Wood at Little London (a Leeds inner-city primary school), with the support of governors and parents, decided that enough is enough and refused to conduct the tests. She is, as we might expect, being pursued in various ways by officials who think school leadership equals blind compliance. This is rather like ordering a bus driver to drive a vehicle with three bald tyres and an engine leaking toxic fumes.

Exploring Alternatives

On 30 March 2017, *More Than a Score* held a seminar at Oxford University to consider alternatives to current modes of assessment, bringing together 35 experts from universities, curriculum associations and the NUT. Presentations covered a range of valuable forms of assessment which have been eclipsed by high-stakes testing, and extracts from printed papers were also circulated and discussed. Many of these are brought together in this current issue of *FORUM*, and a book will follow.

The discussion was both practical and philosophical. It covered concerns about curriculum narrowing, and also the damage to children who were themselves being narrowed as human beings. Schools were becoming exam factories, and children were being 'datafied' – turned into receptacles of added value. As Ruskin argued in the nineteenth century, at the height of the Industrial Revolution:

> The great cry that rises from all our manufacturing cities, louder than their furnace blast, is all in very deed for this – that we manufacture everything there except men; we blanch cotton, and strengthen steel, and refine sugar, and shape pottery; but to brighten, to strengthen, to refine or to form a living human spirit, never enters into our estimate of advantages.

Marx, around the same time, described how processes of alienated labour in factories were estranging human beings from their own humanity: in such a system of production, the worker becomes 'a being with neither needs nor senses ... and from being a man becomes an abstract activity and a stomach'. We can recognise similar processes in a curriculum geared towards high-stakes testing and which neglects so many aspects of the child's development, including practical activity, the creative arts, spoken language and critical thinking.

The first set of articles considers as a whole what primary school assessment should be for, and how it might operate. The set begins with More Than a Score's agreed declaration, 'Assessment: the alternative', which looks at the various purposes served by assessment, and argues that no single form of testing can reasonably perform all these functions. John Richmond's contribution is a carefully constructed proposal for making assessment of spoken language, reading and writing fit for purpose. The extract from Terry Wrigley's submission to the Select Committee entitled 'Some Modest Proposals' presents a critique of the 2016 tests, followed by some proposals for alternatives which would preserve curriculum breadth and encourage more authentic forms of learning.

Denmark is sometimes cited as a country which tests as frequently as England. To dispel some misunderstandings, we include here Jakob Wandall's description 'National Tests in Denmark: CAT as a pedagogic tool' which explains how his country's computer-based tests are designed to avoid damage

to children's wellbeing and education. The author is currently acting as a consultant to the Welsh Assembly for its curriculum reform.

Some years ago Michael Fielding wrote a thought-provoking article, 'Ofsted, Inspection and the Betrayal of Democracy', and around the same time Fred Inglis applied similar thinking to the cult of management, 'A Malediction Upon Management'. Extracts are reprinted here because these ideas also apply strongly to high-stakes assessment.

Pam Jarvis, an expert on child development, points to the profound damage of schooling which ignores children's developmental needs. In 'Homo Sapiens 1.0' she appeals for more holistic care and explains why we must avoid accelerating and intensifying the pace of learning.

Early education researcher Guy Roberts-Holmes compares baseline testing as an attempt to predict each child's future achievement with the processes of observation and pedagogic documentation refined in Reggio Emilia. He contrasts the reductive and decontextualised 'audit style' of the English baseline tests, designed to categorise children, with Malaguzzi's insistence that the teacher must be *open* to possibilities, potentialities and alternatives. The teacher must 'vary, multiply, intensify, re-invent and re-listen to children's activities, behaviours, words and languages. Support and make use of their interests, their forms of learning, choosing and communicating.' Pedagogic documentation 'centres children's voices' and makes children's learning visible through sharing their creative arts, performance, photography, writing and play. These must also be shared in discussion with the child.

Sharen Kucey and Jim Parsons, from Alberta, deepen our understanding of assessment for learning by considering formative assessment in the light of John Dewey's pedagogical principles. They show how well-conceived assessment processes enrich and empower the child and strengthen the partnership with teachers.

The articles that follow each examine a particular form or instance of current practice, and consider practical alternatives. Jonathan Glazzard explains why the phonics check should only be used judiciously with those children who are at a particular stage of struggling with decoding, and presents a range of other assessment tools which teachers could draw upon for specific purposes in accordance with individual needs.

The Key Stage 2 (KS2) writing assessment from 2016 was based on teachers' judgements about a portfolio of writing redrafted over time. Unfortunately, the Testing Agency had so little trust in teachers' professional abilities that tight sets of criteria were imposed, most of them related to surface features of spelling, grammar or punctuation. John Hodgson takes us back to the 1980s, when moderated teacher assessment was used for A-level English literature. He explains how the system was designed to build professional skills and encourage students to show authentic personal responses to their reading, and used for validity and reliability with respect to the study of literature.

Another short piece by John Hodgson criticises the separate testing of grammar and false assumptions about language. John Richmond reminds us of

the folly of equating syntactic complexity with effective writing, with a focus on the opening of Dickens' *Bleak House*. Ros Wilson shows how the marking criteria for KS2 writing have undermined creativity and resulted in artificial and stereotyped writing; in effect, the writing assessment was hijacked to provide a further grammar, punctuation and spelling test. Nerida Spina points to similar effects in Australia, where the anxieties induced by high-stakes testing (the NAPLAN) result in stereotyped paragraphs and alienated students.

Michael Gove's 2014 revision of the National Curriculum deliberately marginalised spoken English, not to mention the children's other languages. It restricted spoken English to formal performances such as reading aloud an extract from a story in primary schools, and in secondary schools it removed the assessment of speaking from the GCSE at age 16. There are deep political implications here, and significantly 16-year-olds attending private schools continue to sit an International GCSE which fully accredits the spoken language as well as written coursework. Valerie Coultas reminds us in this article of the many contributions of speaking and listening across the curriculum, and of the important ways in which we could use spoken language to enrich assessment. Gawain Little, Jo Horn and Steph Gilroy-Lowe also take issue with the narrowness of current assessment modes, with an article focused on mathematics in the primary curriculum. With practical examples they demonstrate the value of observation in the assessment process and argue that this offers a much more effective way of assessing what students know and have learned.

Tony Eaude has just completed research on the use of portfolios to track pupils' learning in the humanities. Taking a broad view of the 'humanities' to include drama and philosophy, for example, and the breadth of spiritual, moral, social and cultural development, he points to a form of assessment which is more holistic and does not disconnect cognitive from ethical and aesthetic aspects of knowledge.

The National Curriculum for England, in its 2014 reform, thoroughly marginalised creativity, both in terms of learners' approaches to 'core subjects' such as English or mathematics, and in terms of the marginalisation of the creative and performing arts. This has been reinforced by modes of assessment which turn writing into a purely technical exercise or insist on one single approved way of solving a maths problem. Two extracts from the late Grant Wiggins demonstrate ways in which different modes of assessment can value and validate creativity. In the first he argues that creativity requires a combination of novelty with appropriateness, and that this is as true in art, philosophy, engineering or athletics as it is in telling jokes. The art of teaching is to design challenges that demand thoughtful problem-solving. He proposes forms of assessment based on descriptive feedback, which makes sense in terms of the task in hand, and particularly in terms of responses from a real audience or a user of a product. The extracts that follow are an attempt at a general rubric for assessing creativity which serves aspirationally to give learners a sense of direction in how to improve.

Terry Wrigley is also concerned with authentic assessment, and represents here the rich task approach developed in Queensland. Rich tasks have to be meaningful and worthwhile in themselves; they draw on a range of skills and knowledge, and culminate in a product or performance. The illustrations provided show clearly the rigour and intellectual challenge of this approach.

The final two contributions return us to the whole structure of assessment in primary schools. Gemma Moss, BERA president at the time of writing this, questions whether the government response to the primary assessment crisis sufficiently recognises the degree of change we need. She questions long-standing assumptions that children make smooth linear measurable progress, that government test instruments are reliable or valid, and that progress data are a reflection of a school's in/effectiveness.

Finally, John Coe of the National Association for Primary Education recalls our attention to the grassroots, especially children and their families, but also their teachers. The current assessment regime works against them, not for them. John speaks of how *they* are affected: 'the children whose study of history is reduced to 30 minutes a week, the parents who deal with the crying child at bedtime because the weekly spelling test is tomorrow'. The parents' revolt in 2016 marks a turning point in resistance which must be built upon. Parents do need sound information, but this can best be provided by teachers coming together to moderate pupils' work, and through narrative reports, not numerical scores. 'For reports to be meaningful they should summarise children's development in all aspects of their growth... More Than a Score applauds schools which produce rich, detailed narrative reports on pupil progress.' The Early Years Foundation Stage Profile could serve as a model.

Sources and Resources

More Than a Score: https://morethanascore.co.uk (with links to video on its own YouTube channel).

Reclaiming Schools: www.reclaimingschools.org, including a regular blog (see 10 October 2016 for links to posts on primary tests).

'Exam Factories? The Impact of Accountability Measures on Children and Young People' (M. Hutchings): https://www.teachers.org.uk/files/exam-factories.pdf

'The Introduction of Reception Baseline Assessment: 'they are children, not robots, not machines' (Alice Bradbury and Guy Roberts-Holmes) https://www.teachers.org.uk/sites/default/files2014/baseline-assessment–final-10404.pdf

'The Mismeasurement of Learning: how tests are damaging children and primary education' (Reclaiming Schools network) https://reclaimingschools.files.wordpress.com/2016/11/mismeasurement.pdf

TERRY WRIGLEY is editor of the international journal *Improving Schools* and Visiting Professor at Northumbria University. He is one of the coordinators of the Reclaiming Schools network (www.reclaimingschools.org). He has a keen interest in curriculum and social justice. His most recent book is *Living on the Edge: rethinking poverty, class and schooling* (co-author John Smyth). *Correspondence*: terrywrigley@gmail.com