Life After Levels

One School's Story of Transforming Primary Assessment

Sam Hunter

00_Hunter_Prelims.indd 3 7/22/2016 4:27:33 PM

Learning Matters
An imprint of SAGE Publications Ltd
1 Oliver's Yard
55 City Road
London EC1Y 1SP

SAGE Publications Inc.
2455 Teller Road
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road
New Delhi 110 044

SAGE Publications Asia-Pacific Pte Ltd
3 Church Street
#10-04 Samsung Hub
Singapore 049483

Editor: Amy Thornton
Development Editor: Jennifer Clark
Production Controller: Chris Marke
Project Management: Deer Park Productions
Marketing Manager: Lorna Patkai
Cover design: Wendy Scott
Typeset by: C&M Digitals (P) Ltd, Chennai, India
Printed and bound by CPI Group (UK) Ltd,
Croydon, CR0 4YY

© 2016 Sam Hunter

First published in 2016

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may be reproduced, stored or transmitted in any form, or by any means, only with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

Library of Congress Control Number: 2016940061

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-1-4739-6426-6 paperback
ISBN 978-1-4739-6425-9 hardback

At SAGE we take sustainability seriously. Most of our products are printed in the UK using FSC papers and boards. When we print overseas we ensure sustainable papers are used as measured by the PREPS grading system. We undertake an annual audit to monitor our sustainability.

1 ASSESSMENT: WHERE WE ARE NOW (AND HOW WE GOT HERE)

This chapter will:

- consider the current 'situation' of assessment in schools;
- explore recent history of assessment;
- examine where we are now.

INTRODUCTION

Theodore Roosevelt believed that 'the more you know about the past, the better you are prepared for the future'. For this chapter I wish to take the former American president's advice and look at the history of assessment in our schools in this country to help us have a better understanding of where we are with assessment now. When did formal assessment start and in what form? What decisions were made by past politicians and educationalists, and with what motives, that have led us down certain paths over the past centuries? How has assessment been viewed over time and has this developed and progressed or does it remain fundamentally the same in its key principles?

TRAINING TO TEACH

When I was training to teach I took the one-year PGCE course following my BA in English Literature. Once you had taken out the holiday breaks and the time spent on teaching practice that left only about 30 weeks (or just 180 days!) of university time to tackle all aspects of teaching and learning: a tough job for any Initial Teacher Training course (ITT). Therefore, it is unsurprising that I was not given an in-depth study of the history of education; it was much more about the here and now and what I needed to teach the children in my future classes to enable them to learn. So I have loved the time I have spent researching this chapter. Sometimes I have got frustrated, while at other times have been excited, as I have read of the trailblazers who have come before us.

01_Hunter_Ch_01_Part_I.indd 5 7/22/2016 2:44:32 PM

IN THE BEGINNING

Assessment in its broadest sense has been in existence from the very beginning. Stone Age man would have assessed if their shooting range was close enough to allow them to kill a wolf while remaining uneaten themselves, while Roman gladiators would have made a judgement on how skilled their opponent was and how likely it was they were going to win the fight. Making accurate assessments helps to keep us safe, from deciding if it is safe to cross the road to testing our food to see if it is cooked properly. Patricia Broadfoot states that, 'Passing judgement is a central part of social behaviour,' and that we are probably unaware of the number of judgements that we make (Broadfoot, 1979, p. 12). You would have made one when you looked at the cover of this book!

ASSESSMENT IN EDUCATION

However, assessment in education took that idea of making a judgement, or an assessment, and then using it to decide on the suitability or success of a learner. History tells us that the record for the first written exam took place at Oxford in 1702 (Black, 1998, p. 10). However, the 'test trail' then goes fairly cold until the nineteenth century when the education system that we recognise today really started to emerge.

It is impossible to separate the history of education from the social history of Britain. Before the 1800s, the profession or job that one held was pretty much determined at birth, due to the circumstances of rank, class or gender. If you were a girl then your education only required skills to help you run a household. If as a son your father was a blacksmith or miner then you were almost certainly going to follow that same path. Likewise, if you were the eldest-born son in a wealthy family then a career in the law or politics was likely while younger brothers might join the army or the clergy. Position and positioning was everything. Therefore, schooling was of relatively low importance. It was not going to help you determine and then facilitate your career choice. What is more, a period of residence of studying, such as four years at university, equalled the qualification; it was the quantity of time rather than the quality of learning that was required to endorse you for your future career (Broadfoot, 1979, p. 29).

I am sure that if I visited my GP and they told me not to worry, that they had spent the appropriate number of years at university but they just hadn't undertaken any assessments in that time I would not be feeling confident! So it was that in 1815 the Society of Apothecaries created a Court of Examiners to examine and award licences and register successful candidates to practise as an Apothecary in England and Wales (Broadfoot, 1979, p. 30). This was a key development in the journey of formative summative assessment and other professions started to follow suit. No longer was it okay simply to have completed the course, but there was now a need to prove that you were competent in what you could do. It also promoted competition for jobs. The assessment system as we recognise it today had been born.

01_Hunter_Ch_01_Part_I.indd 6 7/22/2016 2:44:32 PM

A NEW VIEW OF EDUCATION

There was now the option to study your way out of your social position and as a result society became more mobile as a wider range of professions became increasingly available to a wider range of individuals. Education was therefore viewed in a new way. It became something worth considering as now it really had the power to facilitate a change in one's prospects. From 1833, the government even offered grants to enable poorer children to attend school, widening the reach of opportunity further still. Indeed, nearly all children had some degree of schooling at this time. It was the Newcastle Commission of 1861 which put the cat amongst the pigeons when it sought to review the quality of education that was being provided:

> *We have seen overwhelming evidence from Her Majesty's Inspectors, to the effect that not more than one fourth of the children receive a good education. So great a failure in the teaching demanded the closest investigation; and as the result of it we have been obliged to come to the conclusion that the instruction given is commonly both too ambitious and too superficial in its character … and that it often omits to secure a thorough grounding in the simplest but most essential parts of instruction.*
>
> (Newcastle Report, 1861)

Their solution to improving the quality of education available was to introduce testing with the incentive for teachers being that they would get paid according to pupils' results. No doubt this was viewed as a win-win situation as it would help with relieving the burden of the government budget in supporting schooling for so many children while simultaneously improving standards. However, as Broadfoot comments, this simply led to drilling and rote-learning and frequent testing in the three Rs due to the 'high stake' nature of the tests (Broadfoot, 1996, p. 201). I have to say it feels as though we have made a close return to this with the performance-related pay that is currently finding favour in our present system, where formal summative assessments are perceived to reflect the quality of the teaching. It is also an early indication of the power of tests to restrict the curriculum delivered.

ASSESSMENT AND INSPECTION

Although the first Her Majesty's Inspectors (HMIs) were appointed in 1840, it wasn't until 1899 that the first Board of Education was established. They issued a report in 1911 asking for a review of the role of examinations in their usefulness for preparing school-leavers to go on to employment. However, this was not acted on and instead by 1917 the School Certificate (SC) had been born. As time moved on so too did the expectation placed on achieving end-of-schooling qualifications. Because it was only the grammar schools that could issue the SC, primaries were under greater pressure to ensure that their pupils passed the eleven-plus.

01_Hunter_Ch_01_Part_I.indd 7 7/22/2016 2:44:32 PM

Gipps and Stobart identify the era of the eleven-plus as the 'heyday of the standardised test in primary schools', as schools' success rates were measured by the proportion of pupils going on to grammar schools (Gipps and Stobart, 1993, p. 64). This feels very much like a forerunner to performance tables today, where assessment is used to judge 'success' rather than to review what a child knows.

INTO THE FUTURE

Following the 1944 Education Act, driven by the then Education Minister Richard Butler, all children had been assured a free secondary school place and the chance to gain a qualification, following an increase in the school leaving age to 15. (It would not be until 1972 that this would reach 16.) However, the 1960s and 1970s saw a two-tier education system in place with grammar and private schools each running an alternative education as there were still different qualifications awarded at the end of each route. This eventually changed in 1988 with the arrival of the new GCSE (for the first time, all pupils would be assessed against the same norm-referenced criteria) and the new National Curriculum with its identified key stages for assessment. This is a key moment for us in our whistle-stop tour of the history of assessment for it was at this point that the idea of accountability, reliability and validity of assessments really started to take off. Cue the arrival of: Task Group on Assessment and Testing (TGAT) 1987; 1988 Education Reform Act; Schools Examinations and Assessment Council (SEAC) 1988; Qualification and Curriculum Authority (QCA) created 1997; National Assessment Agency 2004; Ofqual 2008; Standards and Testing Agency (STA) 2011. And so it goes on.

For primary schools, the arrival of the new National Curriculum really was the start of a new era. From 1991, all Year 2 children were assessed at the end of Key Stage 1 with the same hurdles waiting for Year 6 children at the end of Key Stage 2 from 1995. There will be plenty of time throughout this book to discuss the helpfulness, or not, of these particular summative assessments and I shall therefore attempt to avoid getting on my political soapbox too much in the first chapter! However, I feel that I would be sharing the voice of many professionals if I said that the years from 1988 to 2014 were a time of ever-shifting sands. As governments came and went, targets around expected levels of achievement of these assessments continued to change. So too did the format of the tests. I have listed a few key changes below.

- 1990 Key Stage 1 test pilot
- 1991 All Year 2 pupils tested
- 1993 Lord Dearing Review slims down National Curriculum
- 1995 Key Stage 2 tests introduced
- 1996 First primary school/Key Stage 2 performance table published

01_Hunter_Ch_01_Part_I.indd 8 7/22/2016 2:44:32 PM

- 1997 New Labour government sets target of 80 per cent L4+ for Key Stage 2 by 2002
- 2000 Revised National Curriculum
- 2001 New target set of 85 per cent L4+ for Key Stage 2 by 2006
- 2003 Progress measures start to appear on Key Stage 2 performance tables
- 2005 Key Stage 1 external tests are replaced with internal teacher assessment
- 2008 Data for pupils achieving Level 4+ in English AND maths also published
- 2011 Phonics screening pilot
- 2012 First year of phonics screening for all children
- 2013 Introduction of separate Key Stage 2 test for Punctuation, Grammar and Spelling
- 2013 Floor standard that 60 per cent of pupils in a school must get L4+ in reading, writing and maths
- 2014 New National Curriculum
- 2014 Calculator no longer allowed in Key Stage 2 maths test
- 2014 Key Stage 2 Reading test is ramped rather than themed
- 2014 First stage of moving away from assessment through levels
- 2016 First round of testing against the new national standards using an interim framework

These changes are reflected in the 2009 commentary from The Teaching and Learning Programme (TLRP):

> *If policy changes has been hyperactive, in one sense this is not surprising because assessment has been asked to perform an increasing number of functions in recent years; from judging individual pupils to evaluating schools and monitoring national performance.*
>
> (TLRP)

SWIFT CHANGES

For me, the 1990s represent a time of listening to Take That, watching *E.R.* and burning my mouth eating Pop Tarts. Yet in the world of assessment it was the decade that saw the rise of 'Levelling' and then 'Sub-Levelling'. (I need to say here that I gave much thought as to whether these terms required the status of a proper noun and decided that due to the enormity

01_Hunter_Ch_01_Part_I.indd 9 7/22/2016 2:44:32 PM

of their impact that they should!) The initial plan from the Department of Education and Science was to just use Levels. These were broad statements that summarised a level of performance, scaling from Level 1 to Level 8 and beyond. However, as the decade progressed, the focus on data as a means to measure pupil, and therefore school, performance increased. So too did the need to demonstrate progress over shorter periods of time rather than just the end of key stages. First there was the need to make a summative judgement at the end of each school year, and then at the end of each term, or even half-term. The breadth of the Level Descriptors was therefore deemed to be insufficient to show progress over shorter periods of time and in 2001 Average Point Scores (APS) arrived to enable further numerical data-crunching to take place with Sub-Levels following hot on their heels.

I think that it is important to state here that at the point that the National Curriculum and Key Stages and Level Descriptors were created in 1988, Sub-Levels were never part of the plan. At no point did any government or education department give explicit guidance to educators as to what a child was required to do to secure, for example, 3a in reading as opposed to 3b, in the same way that they had for whole Levels. The closest they have come to this were the 'recommended' Assessing Pupil Progress grids (APP), linked to the National Strategies in 2008, which gave instruction towards a best fit judgement, or the thresholds given for Key Stage 2 tests which informed teachers if a child's score enabled them to qualify for a 4b or just a 4c. Teachers were therefore left to make their own judgements as to how far along a Level descriptor a particular child had got at a given point in time. Despite in-school moderation, and external moderation around Years 2 and 6, there was still room for personal interpretation as to what constituted these Sub-Levels. This seems incredible given the importance that APS were given in evaluating school performance.

League tables

School performance became very much a public affair during the 1990s. The first primary league tables were published in 1996. There have been various arguments for why this was a productive move; some have argued that it enabled parents to be fully informed about the choice of where they should send their children to school while others argued that there should be greater accountability for how public money was being spent. There have also been those voices championing the test with public results as the way to drive up standards. However, critics claimed that it created a 'name and shame' culture and, as it was based purely on test scores, was not an accurate indicator of the quality of education provided. Neither did it take into account the context of the school or the starting points of the children. Caroline Gipps states that educational assessment is not high stakes but public test data is (Gipps, 1994, p. 161). League tables remain part of the education landscape today. Some of the concerns stated above have been addressed and progress now carries as much gravitas as attainment, and context of the learners carries more weighting.

01_Hunter_Ch_01_Part_I.indd 10 7/22/2016 2:44:32 PM

However, the publishing of performance tables reinforces the 'high-stake' nature of the end of key stage summative assessments. Coupled with this is the presence of the data analysis produced by the education department and used by those bodies that make external judgements, such as Ofsted. The current format is Raiseonline, produced by the Department for Education. Until 2015, it made rigorous use of APS to track attainment and progress of pupils' performance against national averages. As assessment now moves into the new post-Levels era, the format of this annual data report will need to metamorphose into something new, but its message will, no doubt, continue to carry the same status.

The rise of the Sub-Level also gave rise to the education system's infatuation with 'tracking' and the noughties saw an explosion in paper and computer-based systems for tracking pupil progress. Suddenly children were being coloured-coded into red, amber and green, depending on how close they were to the Sub-Level targets that had been set for them. Schools agonised over how many Sub-Levels were required for 'acceptable' progress over the course of any one academic year. 'Assessment Leader' became a new role, which often meant 'the person who looks at the tracking and spots patterns and identifies children who appear to be falling behind'. Don't get me wrong – the use of tracking data and its analysis can be really helpful for spotting trends and ensuring that children do not get left behind. However, the fixation on tracking can distract us from the process of assessment. You will be able to read more about that in Chapter 3.

However, at the same time as the historical narrative detailed in the pages above was playing out, so too was an alternative one. Remember the 1911 review by the Board of Education? One of their recommendations was to move away from a simple end of phase exam and to look instead at other ways to improve learners' chances as they moved into places of work, such as teachers talking to future employers about pupils' strengths and areas for development (Broadfoot, 1979, p. 33). Then there was the Hadow Report of 1933 which argued against set standards being expected for infant schools: 'In none of this should a uniform standard to be reached by all children be expected. The infant school has no business with uniform standards of attainment' (Swaffield, 2008, p. 10). I think it is also important to mention here too the Plowden Report of 1967 which challenged educators to look at the child as a whole and to centre learning around them. These are examples of past voices who wanted to lead policy down a path away from the dominance of the formal summative assessment.

Added into this are the voices of those outside the government remit who nonetheless sought to affect policy making. The Assessment Reform Group (ARG), formerly known as the Assessment Task Group (ATG), was active from 1989 to 2010. It worked tirelessly to share an alternative view of assessment in the form of formative judgements which impact directly on learning – you will be able to read more about that in later chapters. Paul Black and Dylan Wiliam's key piece from 1990 on formative and summative assessment is now compulsory reading for anyone who is involved in teaching, yet I do not recall *Inside the Black*

01_Hunter_Ch_01_Part_I.indd 11 7/22/2016 2:44:32 PM

Box being on the book list when I was doing my PGCE in 1996/7. How much I missed! Throw into the mix Carol Dweck's pivotal work on growth mindsets and John Hattie's research on what really makes an impact on learning and the stage is set for a battle of minds rather a neat meeting of them.

THE PRESENT DAY

And so we come to the present day. The arrival of the new National Curriculum in 2014 gave the perfect opportunity for a review of the Levels assessment system to see if it remained fit for purpose. The Department for Education stated in May 2013 that as there was to be new curriculum content, a new assessment system needed to be established and consequently Levels would be permanently removed. Tim Oates, who chaired the expert panel that reviewed the National Curriculum between 2010 and 2013, explains the reasons for moving away from Levels. First, there was the concern around the negative affect on children's learning when they used Levels to label themselves, for example as Level 3, and how this had the power to place a ceiling on expectation. In addition to this lay the perceived clash between the unnecessary pace that was driving children though Levels and the core principle of the new National Curriculum of studying fewer key concepts in more detail. He states that teaching and learning should focus more on the breadth and mastery of constructs, rather than the need to move constantly on to the next thing. He also talks about the inconsistencies in the use of Levels, presenting three different ways in which Levels were decided upon. Finally, he explains that the top-performing countries in the world do not use a Levels assessment system (Oates, 2014).

So, from September 2014, teachers of children in Years 1, 3, 4 and 5 embarked on a new assessment journey, with those responsible for Years 2 and 6 joining in a year later. The Department for Education was clear that no guidance would be given to schools as to how they were to replace the old system. It was for schools to decide what was best for them and the needs of their new curriculum and their learners. I have heard Mick Walker, former Executive Director of Education at the QCA and advisor to the NAHT's (National Association of Headteachers) commission on assessment without levels, observe that the teaching profession complains when the government tells them what to do but that equally they moan when they are not told what to do. Some local authorities gave direct instruction or 'recommendation' to their schools while others remained silent, possibly biding their time in case future central instruction was upcoming. This was not the case.

In May 2014, nine schools were awarded the Assessment Innovation Award grants, although this is now only recorded as being eight schools. The aim here was to invest in grass roots projects and to share best practice and resources. My own school, Hiltingbury

01_Hunter_Ch_01_Part_I.indd 12 7/22/2016 2:44:32 PM

Junior School, was the ninth school but as we ultimately did not take the grant money we no longer appear on any of the literature. That was the furthest the Department for Education went in making any recommendations. A Commission on Assessment without Levels met from March to May 2015, publishing a final report in September 2015. Chaired by John McIntosh, the expert panel put forward six key recommendations around how the profession can best support one another moving forward without Levels. However, they made it clear that, 'The report does not provide schools with a template for assessment without levels but offers guidance and support to help schools in designing their own assessment policies, in parallel with their curriculum policies' (McIntosh, 2015, p. 4).

FUTURE OUTLOOK

From 2016, pupils' attainment at the end of Key Stages 1 and 2 is instead to be measured with Scaled Scores. Progress measures will remain in a state of flux until 2020 where the Key Stage 2 Scaled Scores will need to be compared against the old Key Stage 1 Average Point Scores. Throw into this mix the latest Year R baseline assessment and the sands appear to set to shift until at least 2023.

Chapter summary

- The basis of our education and assessment system holds its roots in the social history of this country.
- Formal summative assessments came about as a way to ensure quality of learning and suitability to carry out a profession competently.
- Formal summative assessments have fed notions of selection and competition.
- Assessment has been used over time to hold the performance of schools and teachers to account.
- Formal summative assessment has held the lion's share of authority in school accountability over hundreds of years.
- Although the format of formal summative assessments continues to change, their use for judging school performance remains constant.
- There has been a growing voice to champion the role of formative assessment in our education system.
- The latest assessment system has gone the furthest in fusing summative and formative assessments together.

01_Hunter_Ch_01_Part_I.indd 13 7/22/2016 2:44:32 PM

References

Black, P. (1998) *Testing: Friend or Foe? Theory and Practice of Assessment and Testing.* London: Falmer Press.

Broadfoot, P. (1979) *Assessment, Schools and Society.* London: Methuen.

Broadfoot, P. (1996) *Education, Assessment and Society.* Buckingham: Oxford University Press.

Duke of Newcastle (1861) *The Royal Commission on the State of Popular Education in England* (Newcastle Report). http://www.educationengland.org.uk/history/chapter03.html (Accessed 18.2.16)

Gipps, C. (1994) *Assessment: Beyond Testing: Towards a theory of educational assessment.* London: Falmer Press.

Gipps, C. and Stobart, G. (1993) *Assessment: A teachers' guide to the issues.* London: Hodder and Stoughton.

McIntosh, J. (chair) (2015) *Final Report on the Commission on Assessment Without Levels.* London: Department for Education and Standards and Testing Agency.

Oates, T. (2014) *National Curriculum: Tim Oates on assessment.* Department for Education https://www.youtube.com/watch?v=-q5vrBXFpm0 (Accessed 17.1.16)

Swaffield, S. (ed.) (2008) *Unlocking Assessment: Understanding for reflection and application.* London: Routledge.

Teaching and Learning Research Programme (2009) *Assessment in Schools: Fit for purpose?* London: TLRP.

Further reading

Conner, C. (1991) *Assessment and Testing in the Primary School. Basingstoke*: Falmer Press.

Desforges, C. (1989) *Testing and Assessment*. London: Cassell.

Gipps, C. (ed.) (1992) *Developing Assessment for the National Curriculum*. London: Kogan Page.

Gipps, C. and Murphy, P. (1996) *A Fair Test?* Buckingham: Open University Press.

Gipps, C., Steadman, S., Blackstone, T. and Stierer, B. (1983) *Testing Children: Standardised testing in local education authorities and schools*. London: Heinemann.

Horton, T. (ed.) (1990) *Assessment Debates*. Milton Keynes: Open University Press.

Hursh, D. (2008) *High-stakes Testing and the Decline of Teaching and Learning: The real crisis in education.* Plymouth: Rowman and Littlefield.

Norris, N. (1990) *Understanding Educational Evaluation*. London: Kogan Page.

Pole, C. (1993) *Assessing and Recording Achievement.* Buckingham: Open University Press.

Sattery, D. (1989) *Assessment in Schools*. Oxford: Basil Blackwell.

Sumner, R. (1987) *The Role of Testing in Schools*. Windsor: NFER: Nelson.

01_Hunter_Ch_01_Part_I.indd 14 7/22/2016 2:44:32 PM