

The Caroline Benn Memorial Lecture 2008

Testing, Testing, Testing

Lecturer: Dr Bethan Marshall

Tony Benn, who attended with his daughter Melissa spoke briefly of his family's appreciation that the SEA had staged the annual lectures in honour of Caroline and her work.

Professor Richard Pring, as Chair, introduced the speaker.

Dr Marshall, who is particularly involved with the teaching, and testing, of English in schools drew her comments from many sources, including noted educationalists and the results of various academic surveys.

Her starting point was

The Real Cost of SATS 1995:

“The SATs have a negative influence on the curriculum because they narrow and limit what can be done. They tend to eliminate creativity and imagination in both the teachers and the student. Instead we are told what to do, what play to read, and what scenes will be examined.”

and the ATL survey 1996 which found “teaching to the test more than was reasonable” ... ” The tests [had] narrowed the curriculum”

The trouble with tests is not a recent issue:

Matthew Arnold (1822 to 88), poet, critic and teacher:

“By ingenious preparation (as it is now found possible) to get children through the Revised Code examination in reading writing and ciphering, without their really knowing how to read write or cipher, so it will in practice, no doubt be found possible to get three fourths of the one fifth of the children over six through the examination in grammar, geography, and history, without their really knowing any one of these three matters.”

“All test examinations may be said to narrow reading upon a certain given point, and to make it mechanical.”

“It tends to make instruction mechanical and to set a bar to duly extending it ... [and] must inevitably concentrate the teachers' attention on producing this minimum and not simply on the good instruction of the school. The danger to be guarded against is the mistake of assuming these two – the producing of the minimum successfully and the good instruction of the school – as if they were identical.”

More recently:

“Drill the child into the mechanical production of quasi-material results; and his success in doing this will be gauged in due course by an "examination" a periodic test which is designed to measure, not the degree of growth which the child has made, but the industry of the teacher as indicated by the receptivity of the class.” (Holmes).

and Tomlinson:

“ The volume of assessment is significant: This year there were around 7.5 million qualification entries in the national school and college qualifications (GCSEs, VCE and GCE A levels, AEs and GNVQs) and 11.6 million unit entries, covering 262 specifications of learning. Awarding bodies deal with 26 million scripts each year.

All this reflects the fact that most learners are examined at 16, 17 and 18. In addition, doing more subjects at the same level is the main way in which stretch is provided for the most able learners. Our proposals challenge both this definition of stretch and the notion that learners need to be repeatedly assessed through external examinations during the 14-19 phase. Learners would be stretched through broader programmes, more varied learning and assessment styles and the possibility of moving faster to a higher level, rather than simply doing more subjects.”

Limitations of testing

Functional Skills

“The initial stage in the process of defining a qualification . . . for the purposes of ensuring that the content of knowledge, understanding and skills developed is captured. They allow the scope, content, level of demand and parameters of the areas of knowledge, understanding and skills to be defined”.

Reading

‘Read, understand and compare texts and use them to gather information, ideas, arguments and opinions’.

Writing

‘Write documents communicating information, ideas and opinions effectively’

Speaking and listening

‘Make a range of contributions to discussion and make effective presentations’.

A series of evaluations conducted in **Queensland** led to these observations on coursework-based assessment:

Relies heavily on teacher involvement

Course work moderated by schools and then expert teachers

Standardising meetings – constant meeting and discussion of coursework led to a raising of professional standards and an agreement of grades.

The criteria led assessment with ‘respect for teacher professionalism in judging student achievement’ (Cummings and Maxwell, 2004).

From the JMB:

The lively interest which should be aroused by learning to read and to write English is killed, so it is asserted, by the need to prepare for writing answers to stereotyped questions on GCE language Papers. (Wilson, in Hewitt and Gordon, 1965)

It is more likely, however, that internal assessment is proposed because of a conviction that it is a more valid way of assessing the attributes or the skills which are involved. Increased validity could result for two reasons: first because the assessment of skills concerned may be difficult or impossible to achieve by external examination and second, because assessment on a single occasion may be a totally inadequate test of a candidate's overall competence. (Smith, 1978)

Science Education

“ We start out with the intention of making the important measurable, and end up making the measurable important. The effect of this has been to weaken the correlation between standardized test scores and the wider domains for which they are claimed to be an adequate proxy.”(Wiliam, 2006)

At the time of giving the lecture Dr Bethan Marshall was Senior lecturer in English education at King's College, London.