Educational Standards in Historical Perspective

RICHARD ALDRICH

Introduction

THE *Oxford English Dictionary* lists some thirty usages of the word, 'standard'. Two broad categories amongst such usages are those of a military or naval ensign and an exemplar of measure or weight. The term, 'standard', first appears in English with reference to the Battle of the Standard, fought at Northallerton between the English and Scots on 22 August 1138. A contemporary chronicler, Richard of Hexham, described the standard as a mast of a ship surmounted by flags around which the English grouped, which was called a standard because 'it was there that valour took its stand to conquer or die'. In this sense the 'raising of one's standard' meant (and still may imply) setting forth to engage in battle or other stirring deeds.

The second sense, of an authorised exemplar of measure or weight, for example the standard lengths built into the wall of the observatory at Greenwich, is connected to the first inasmuch as early usages state or imply 'the king's standard'. Just as the royal standard in battle was the place around which all should rally and from which commands were issued, so the royal or official standards or measures were those which subjects should employ in business and commercial dealings. This sense of a measure to which all objects or persons should conform was extended into many spheres of life, becoming a definite degree of quality, viewed as a measure of an adequate level for a particular purpose or as a prescribed object of endeavour. In 1862, with the introduction of the Revised Code into elementary schools, the word 'standard' took on a further meaning, defined by the *Oxford English*

Proceedings of the British Academy, **102**, 39–67.

Dictionary as, 'Each of the recognised degrees of proficiency, as tested by examination, according to which school children are classified'. From the end of the nineteenth century, the demise of centrally-controlled annual examinations led to the gradual disappearance of this particular connotation of the word 'standard'. Today, however, standards in education are as prominent an issue as they were a century and more ago. In 1998, a School Standards and Framework Act stands on the Statute Book, a Standards and Effectiveness Unit with more than 100 members of staff has been created within the Department for Education and Employment, while the Standards Task Force is chaired by the Secretary of State, David Blunkett, himself, with Chris Woodhead, the Chief Inspector of Schools, and Tim Brighouse, Director of Education for Birmingham, as vice-chairs.

Dictionary definitions demonstrate that the terms 'standard' and 'standards' have many different meanings. These meanings have changed over time, and will continue to do so. Confusion can, and does, occur, even in official documents. For example, the third section of the White Paper, *Excellence in Schools*, published in 1997, is entitled 'Standards and accountability'. Its first sub-heading declares, 'Raising standards: our top priority', and it reports that 'in the 1996 national tests only 6 in 10 of 11 year-olds reached the standard in mathematics and English expected for their age' (DfEE 1997: 10). Here, the term, standard, is being used in the sense of an accepted level against which all should be judged. But the raising of standards referred to in the sub-heading presumably does not apply (at least not initially) to raising the expected standards which four out of ten children were already failing to reach. Rather it refers to raising the unacceptable standards or levels of achievement of those who were failing to reach the expected standard. Indeed, the White Paper states categorically that by the year 2002

> 80% of 11 year-olds will be reaching the standards expected for their age in English; and 75% of 11 year-olds will be reaching the standards expected for their age in mathematics.
>
> (DfEE 1997: 19)

It remains to be seen whether any modification of the expected standards will be needed to ensure that these percentages are achieved.

Further questions may be identified here, questions which are historical in that they relate to continuities and changes in human purposes and judgements over time. For example, if an educational standard is defined as an authorised exemplar or measure, who defines

that standard? Is it intended to be the same for all pupils of the same age? Are all pupils expected to achieve it, or is its purpose to select some and to reject others—as in the case of the eleven-plus examination which governed transition from the primary stage before the introduction of comprehensive secondary schools, or in the use of GCE Advanced level grades in selection of entrants to higher education. Standards, in common with examinations, curricula and education itself, have a history which has been contested by contemporaries and historians alike. Since 1988, government policies to improve the quality of education have been based upon the concept of an expected standard of achievement for all children of a particular age. This situation may be strongly contrasted with the advice given in the Plowden Report of 1967 which 'concluded that it is not possible to describe a standard of attainment that should be reached by all or most children. Any set standard would seriously limit the bright child and be impossibly high for the dull. What could be achieved in one school might be impossible in another' (DES 1967: i, 201–2).

In further contrast, in 1971 Cox and Dyson included the following ironic 'progressive' definition of standards: 'Irrelevant academic concept designed to exclude, or penalize, students distinguished for *either* concern *or* creativity *or* both' (Cox and Dyson 1971: 215).

Today, it is argued by central government and its agencies not only that it is necessary to define national standards, but also that the levels of educational attainment of many children are lower than they should be because the expectations of many teachers, parents and pupils are too low. Teacher, pupil and parental perceptions of importance and standards, however, may conflict. For example, Schools Council Enquiry I, *Young School Leavers*, published in 1968, showed that parents of 15 year-old leavers, unlike their teachers, placed the greatest emphasis upon doing well in studies that would enable their children to get jobs. This was understandable, given that until very recent times most children in Britain left school at the earliest opportunity and proceeded directly into employment. Similarly, in October 1998, the *Times Educational Supplement* reported that 'Parents do not share ministers' high level of concern about academic achievement in schools' (Dean 1998). Expectations of standards, in common with definitions, therefore, exhibit changes, as well as continuities, over time. Levels of expectations in terms of educational achievement are the product of a long and contested history in which governmental priorities, economic and religious doctrines, employment requirements, and social factors,

including those of class and gender, as well as the expectations of teachers, parents and pupils, have loomed large.

The use of the term, 'historical perspective' in the title of this paper, and the application of historical perspectives to contemporary educational issues (Aldrich 1996), also merit a brief explanation. History may be defined as the disciplined study of human events with particular reference to the dimension of time—principally in the past, but also with some acknowledgement of present and future. Such acknowledgement is essential, if only because that which is now past was once both a future and a present. Contemporary contests around the issue of educational standards frequently draw upon historical perspectives. For example, George Walden, a Conservative education minister, 1985–87 and columnist for the *Daily Telegraph*, has recently traced the perceived low standards of the English education system of today to the long-standing social class divide between private and state schools. Walden argues that while the seven per cent of children in private schools flourish, as indicated by levels of achievement in public examinations at ages 16 and 18, and by the 90 per cent of pupils from private schools who proceed to higher education, the remaining '93 per cent are still locked into a second-class system of education' (Walden 1996: 1). One major reason for this inferiority, according to Walden, is that teachers in private schools 'have remained largely immune to the social dogmas and experimental methods inflicted on generations of state school pupils' (Walden 1996: 44). Other commentators, however, including many teachers in state schools, would lay greater emphasis upon the link between educational standards and financial resources. Figures produced by Walden himself show the annual cost per secondary day pupil in state schools as being £2,250, with a pupil-teacher ratio of 18.4, as opposed to secondary day pupils in private schools with a cost of between £3,600 and £8,700 and a pupil-teacher ratio of 9.8 (Walden 1996: 43).

The journalist, Melanie Phillips, is another contemporary high-profile commentator on educational standards. Her book, *All Must Have Prizes*, also published in 1996, begins with a catalogue of evidence to demonstrate 'Standards sliding' (Phillips 1996: 1–6). Phillips, a columnist for the *Guardian*, *Observer* and *Sunday Times*, declares that today, 'The rot sets in at primary school level and runs throughout the system.' (Phillips 1996: 5) Her key chapter seven, entitled 'The Unravelling of the Culture', is historical. It traces 'the collapse of external authority that lies at the heart of the breakdown in education' (Phillips

1996: 187), from the Enlightenment and Rousseau, through a list of malign progressives which includes Holmes, Dewey, Nunn, Isaacs, Piaget, Simon and Stenhouse.

The belief that standards were better in the past has considerable nostalgic appeal and is frequently urged in the popular press. The following excerpts from the *Daily Mail* are typical: 'the brutal truth is that standards have fallen'; 'most parents and many teachers believe that children are less literate and numerate than they were 20 years ago'. From the opposite end of the political spectrum the *Daily Mirror* assures its readers that: 'literacy in Britain is marching backwards'; 'general educational standards have slipped alarmingly in the past decade or so' (Aldrich 1997: 9–10). These quotations, however, date from 1975 and 1976. Educational standards may, indeed, have fallen in the 1970s as in the 1990s, and in the 1980s as well, but due allowance must be made for the polemical style of many journalists, and for the well-attested fact that bad news sells more copies of newspapers (and of some books) than good.

Changes

Substantial changes in educational standards across time may be simply demonstrated. For example, in Britain, as in the western world in general, there was a general though uneven, rise in literacy levels across several centuries. Significant studies of such phenomena include those by Cipolla (1969), Clanchy (1979), Cressy (1980), Stephens (1987) and Vincent (1989). One important feature of these historical studies is their emphasis upon such factors as occupation and general culture in the increase (and occasional stagnation and decrease) in literacy levels. Thus Cressy notes that stagnation in the development of literacy during the second half of the eighteenth century has been associated by some historians with the 'social disruptions of the industrial revolution' (Cressy 1980: 177). Vincent explores the changing patterns of nineteenth-century male and female literacy, and highlights the contribution to literacy development of the penny post, and such associated features as Christmas and Valentine cards and the picture postcard. In 1858 'the Postmaster General had drawn attention to the fact that as many letters were being delivered in Manchester alone as in the whole of Russia' (Vincent 1989: 46).

Broad comparisons of literacy across centuries and cultures help to confirm the complex definitional problems associated with educational

standards. In medieval Britain, education was construed primarily in vocational and religious contexts. Boys learned the skills of their fathers, and girls of their mothers; all were taught the basic elements of the Christian faith. Teaching and learning were essentially oral. The majority of people had neither the opportunity nor the immediate need to acquire literacy. Literacy, itself, is a term which is as difficult to define as to measure. The two skills of reading and writing have often been quite separate. In the early modern period, following the development of printing in the fifteenth century and the religious Reformation of the sixteenth, increasing numbers of people in Protestant countries learned to read the Bible and other religious works. This did not, however, necessarily mean that they also read secular literature, nor that they learned to write. The nature and extent of literacy (and numeracy) needed for an individual to function effectively in a particular society has clearly changed over time. Current debates about literacy levels should be set in a series of contexts which include the impact of such recent developments as the popular newspaper and typewriter in the nineteenth century, and television and the computer in the twentieth.

There is evidence to indicate that the steady improvement in literacy standards which took place across five centuries in Britain has not been maintained in the second half of the twentieth century. In a recent paper, presented at conferences in 1997, Greg Brooks of the National Foundation for Educational Research argues that during the period from 1948 until 1996 literacy standards in the United Kingdom changed very little. Indeed, there was a slight fall among eight year-olds (children in year 3) in England and Wales during the late 1980s, followed by a recovery in the early 1990s. This fall might have been associated with the introduction of the National Curriculum, which reduced the amount of time devoted to literacy in primary schools, and with the high number of teachers leaving the profession at that time. International evidence suggests that the levels achieved by high and middling performers in the United Kingdom are comparable to the best in the world, although among children and adults there is a significant proportion of the population who have poor literacy skills. Brooks concludes that the most effective way of raising average levels of achievement would be to 'intervene early in the education of children who are already failing or at risk of doing so, to ensure that they are equipped with the literacy (and numeracy) skills necessary for the rest of their education and for life' (Brooks 1997: 1).

First reports of a follow-up study to the ORACLE (Observational,

Research and Classroom Learning Evaluation) project, carried out in sixty East Midlands primary classrooms between 1976 and 1978, suggest not only stagnation, but actual decline. In July 1998, Maurice Galton reported in the *Times Educational Supplement* that his comparisons of children in years 4, 5 and 6 at the end of the school years 1976–7 and 1996–7 showed significant decline in the three basic areas of mathematics, reading and language skills. Galton judged that

> The fall appears to have occurred in the late 1980s and throughout the 1990s. The one factor which stands out in this period of rapid change is the national curriculum . . . Teachers said they were under pressure to get through the curriculum, emphasising instruction and content rather than teaching for understanding. . . . Teachers told us, despite denials from the Office for Standards in Education, that it is easier to pass inspections if you have a secondary-style timetable to demonstrate that the requisite hours are given to the core subjects. . . . It is perhaps ironic that those who have criticised primary teaching most vehemently, such as the Chief Inspector, helped to encourage this form of the national curriculum. (Galton 1998)

This evidence of stagnation or decline at primary levels in the 1980s and 1990s must be set against other evidence from the secondary, further and higher education sectors. Although commentators such as Phillips cite evidence of decline at all levels, there can be no doubt that over the same period there has been a steady increase in pupils achieving passes in public examinations at ages 16 and 18, while the numbers entering higher education and attaining degrees have more than doubled.

The Revised Code

In 1858 the Newcastle Commission was appointed to examine the condition of popular education in England. It reported in 1861. The Revd. James Fraser, an assistant commissioner, investigated elementary schooling in Devon, Dorset, Somerset, Herefordshire and Worcester. This substantial extract from his report indicates the religious, occupational, social class and gender contexts in which the most contentious example of government-directed attempts to raise educational standards took place.

> Even if it were possible, I doubt whether it would be desirable, with a view to the real interests of the peasant boy, to keep him at school till he was 14 or 15 years of age. But it is not possible. We must make up our minds to see the last of him, as far as the day school is concerned at 10 or 11. We must frame our system of education upon this hypothesis; and I venture to maintain that it is quite possible to teach a child soundly and thoroughly, in a way that he shall

Research and Classroom Learning Evaluation) project, carried out in sixty East Midlands primary classrooms between 1976 and 1978, suggest not only stagnation, but actual decline. In July 1998, Maurice Galton reported in the *Times Educational Supplement* that his comparisons of children in years 4, 5 and 6 at the end of the school years 1976–7 and 1996–7 showed significant decline in the three basic areas of mathematics, reading and language skills. Galton judged that

> The fall appears to have occurred in the late 1980s and throughout the 1990s. The one factor which stands out in this period of rapid change is the national curriculum . . . Teachers said they were under pressure to get through the curriculum, emphasising instruction and content rather than teaching for understanding. . . . Teachers told us, despite denials from the Office for Standards in Education, that it is easier to pass inspections if you have a secondary-style timetable to demonstrate that the requisite hours are given to the core subjects. . . . It is perhaps ironic that those who have criticised primary teaching most vehemently, such as the Chief Inspector, helped to encourage this form of the national curriculum. (Galton 1998)

This evidence of stagnation or decline at primary levels in the 1980s and 1990s must be set against other evidence from the secondary, further and higher education sectors. Although commentators such as Phillips cite evidence of decline at all levels, there can be no doubt that over the same period there has been a steady increase in pupils achieving passes in public examinations at ages 16 and 18, while the numbers entering higher education and attaining degrees have more than doubled.

The Revised Code

In 1858 the Newcastle Commission was appointed to examine the condition of popular education in England. It reported in 1861. The Revd. James Fraser, an assistant commissioner, investigated elementary schooling in Devon, Dorset, Somerset, Herefordshire and Worcester. This substantial extract from his report indicates the religious, occupational, social class and gender contexts in which the most contentious example of government-directed attempts to raise educational standards took place.

> Even if it were possible, I doubt whether it would be desirable, with a view to the real interests of the peasant boy, to keep him at school till he was 14 or 15 years of age. But it is not possible. We must make up our minds to see the last of him, as far as the day school is concerned at 10 or 11. We must frame our system of education upon this hypothesis; and I venture to maintain that it is quite possible to teach a child soundly and thoroughly, in a way that he shall

> not forget it, all that is necessary for him to possess in the shape of intellectual attainments by the time that he is 10 years old. If he has been properly looked after in the lower classes, he shall be able to spell correctly the words that he will ordinarily have to use; he shall read a common narrative—the paragraph in the newspaper that he cares to read—with sufficient ease to be a pleasure to himself and to convey information to listeners; if gone to live at a distance from home, he shall write his mother a letter that shall be both legible and intelligible; he knows enough of ciphering to make out, to test the correctness of, a common shop bill; if he hears talk of foreign countries he has some notions as to the part of the habitable globe in which they lie; and underlying all, and not without its influence, I trust, upon his life and conversation, he has acquaintance enough with the Holy Scriptures to follow the allusions and the arguments of a plain Saxon sermon, and a sufficient recollection of the truths taught him in his catechism, to know what are the duties required of him towards his Maker and his fellow man. I have no brighter view of the future or the possibilities of an English elementary education, floating before my eyes than this. (Newcastle Report 1861: XXI (ii) 46–7)

Not that Fraser believed that one half, or even a quarter of children who left school aged 10 did 'carry with them into the business of life even the humble amount of accomplishments which I have named. But they ought to do; and in all the schools which in my list (Table XVIII) I have named as "efficient" I believe they do.' (Newcastle Report 1861: XXI (ii) 47)

Under the Revised Code of 1862 a large part of the central government's financial assistance to aided elementary schools was based upon the principle of payment by results. Annual examinations were carried out by Her Majesty's Inspectors, who were issued with detailed instructions for the purpose. Reading, and the slate work of younger children in writing and arithmetic, were to be examined in the school. The paper work of older scholars might be marked in the school, but all work done on paper, together with the mark schedule, had to be sent to the Education Department. The six standards established in 1862, which roughly corresponded to children aged between six and 12 are shown in Table 1.

A major Code revision occurred in 1872. The original Standard I was abolished; the first examination of children would now normally begin at age seven. The existing Standards II to VI were re-numbered I to V. A new Standard VI was added, as shown in Table 2.

At the same time it was announced that henceforth no pupil could be presented for examination for a second time under a lower standard or for the same standard. Additionally, from 31 March 1873, no day

Table 1. Standards I–VI as established in 1862.

	Reading	Writing	Arithmetic
Standard I	Narrative in monosyllables.	Form on blackboard or slate from dictation, letters capital and small manuscript.	Form on blackboard or slate, from dictation, figures up to 20: name at sight figures up to 20: add and subtract figures up to 10, orally and from examples on the blackboard.
Standard II	One of the narratives next in order after monosyllables in an elementary reading book used in the school.	Copy in manuscript character a line of print.	A sum in simple addition and subtraction and the multiplication table.
Standard III	A short paragraph from an elementary reading book used in the school.	A sentence from the same paragraph, slowly read once and then dictated in single words.	A sum in any simple rule as far as short division (inclusive).
Standard IV	A short paragraph from a more advanced reading book used in the school.	A sentence slowly dictated once, by a few words at a time from the same book, but not from the paragraph read.	A sum in compound rules (money).
Standard V	A few lines of poetry from a reading book used in the first class in the school.	A sentence slowly dictated once, by a few words at a time, from a reading book used in the first class of the school.	A sum in compound rules (common weights and measures).
Standard VI	A short ordinary paragraph in a newspaper, or other modern narrative.	Another short ordinary paragraph in a newspaper, or other modern narrative, slowly dictated once by a few words at a time.	A sum in practice or bills of parcels.

Table 2. Standard VI from 1872.

Standard VI	To read with fluency and expression.	A short theme or letter, or an easy paraphrase.	Proportion and fractions (vulgar and decimal).

scholar over nine years of age and no evening scholar over 13 could be presented in Standard I, while from 31 March 1874 no day scholar over nine years of age, and no evening scholar above 14 could be presented for Standard II (CCE 1872: lxxxiii).

The results for England and Wales in the year ending 31 August 1872, the first year of the operation of the New Code, are shown in Table 3 (CCE 1873: xi-xii).

Comparison of standards during the period of payment by results is difficult. The addition of 'specific' subjects, which were examined individually, and 'class' subjects in which the overall proficiency of the class was assessed, provided other ways of securing grants. For example, in 1872 of the 118,799 children presented in Standards IV–VI, 71,507 were also examined in one or more of the specific subjects. Of these 49,273 secured passes, of whom 18,958 did so in two subjects, with geography, grammar and English history proving to be the most popular. The most obvious feature, however, was the large increase in the numbers of

Table 3. Examination results for England and Wales in the year ending 31 August 1872.

Number of day scholars qualified for examination	792,706
Number presented for examination	661,589
Number presented in Standard I	258,946
Number presented in Standard II	172,391
Number presented in Standard III	111,453
Number presented in Standard IV	66,925
Number presented in Standard V	36,843
Number presented in Standard VI	15,031
Number examined under 10 years of age	342,655
Number examined over 10 years of age	318,934
Number examined in Standards I–III	
Under 10 years of age	339,618
Over 10 years of age	203,172
Number examined in Standards IV–VI	
Under 10 years of age	3,037
Over 10 years of age	115,762
Number who passed without failure in any subject	
Standards I–III	
Under 10 years of age	213,395
Over 10 years of age	122,704
Standards IV–VI	
Under 10 years of age	1,814
Over 10 years of age	63,982

schools seeking government recognition as public elementary schools for the purposes of obtaining grants. The New Code was introduced in the immediate aftermath of the Elementary Education Act of 1870 which established school boards empowered to levy a rate for education. In consequence, whereas in the period 1862 to 1869 the average number of schools seeking inspection for the purposes of obtaining grants was 492, in 1870 the figure was 1,114 and in 1872, 1,530. The annual report of the Committee of Council on Education for 1872–3 welcomed the increase in applications, and thus in scholars presented for examination. But it also regretted the considerable numbers of scholars who were not entered for any examinations, the great preponderance of scholars in Standards I–III as opposed to IV–VI, and the very high percentage of those aged 10 or over who were presented in Standards I–III.

Joseph Payne, a former schoolmaster who in 1873 became first professor of the College of Preceptors, was one of the major critics of standards under the system of payment by results (Aldrich 1995: 179–85). In September 1872 he presented a paper entitled 'Why are the results of our primary instruction so unsatisfactory?' to the annual congress of the Social Science Association in Plymouth. The tone of his paper was calculated to shock: 'This, then, is the final result of the working of 15,000 schools, conducted by 26,000 teachers, at a cost of about one million a year. All this stupendous machinery is contrived and kept in motion to send out into the world annually about 16,000 children with the ability to read, write and cipher moderately well' (Payne 1872: 245).

Another opponent was John Menet, Vicar of Hockerill, whose 26-page pamphlet, *A Letter to a Friend on the Standards of the New Code of the Education Department*, was published in 1874. In this work Menet posed the question 'Why, then, is the principle of Standards radically bad?' (Menet 1874: 7) and grouped his answers under eight main themes.

The first was that it was impossible to devise standards for the whole country. No two schools were alike. Some were established and settled, others were new and reliant on shifting populations. Some were large with children classified according to standards, others had all pupils in the one class. Some children walked to school along a few yards of pavement, others trudged for hours across muddy field, moorland or fen. Given the variety of school circumstances it was obvious that a single set of standards would be much too difficult for some and much too easy for others.

The second problem was that the minimum requirements laid down by the standards too often became the maximum. If a penalty was imposed for not reaching a particular level, then the chief aim was to avoid the penalty, and not to aspire to anything much beyond it. The grudging spirit in which the Code was operated was indicated by the instructions to inspectors issued from the Council Office in September 1862. These advised that inspectors must be satisfied as before on a range of matters—including the state of the buildings, qualifications of the teachers, and keeping of registers—and that the new examination would not supersede judgments in these matters but rather presuppose them. The examination results did not prescribe that '*if thus much is done, a grant shall be paid*, but, *unless thus much is done, no grant shall be paid*' (CCE 1863: xviii).

Menet's third point was the effect of standards on the quicker and slower children. Quicker children were held back, for there was a serious disadvantage to the school in children passing through the standards too quickly. One standard per year was the most they should progress in order to get the maximum grant. On the other hand, the slowest children could not keep up with the rate of one standard per year. It was for this reason that so many children were presented in the lower standards, and some not at all. From a financial point of view, schools had every inducement to neglect children who had little or no chance of being successful.

His fourth criticism, and that which was most frequently made against the system, was that standards encouraged a dull and mechanical routine of teaching and learning. In 1872 Payne had complained that the system of payment by results was 'mechanical in conception, mechanical in means, mechanical in results . . . Making quantity not quality the test of your results, you shall fail in securing either quantity or quality. The experiment which has now been tried for ten years in England ought henceforth to take a place in the annals of education as an example to deter.' (Payne 1872: 247–8) As HMI Kennedy acknowledged in his report for 1872, under the aegis of standards, most elementary schools merely became glorified infant schools in which nothing but reading, writing and arithmetic were taught at increasing levels of complexity. Such a curriculum stifled true mental progress and development. The teacher's task became to secure as many passes as possible. One of the most telling sections in Menet's pamphlet was his use of quotations from such experienced inspectors as Arnold, Campbell, Fussell, Kennedy, Mitchell, Stewart and Watkins, all of

whom contrasted the superior teaching and learning that went on in many schools before the introduction of the Revised Code. For example, Matthew Arnold was quoted from his report for 1872–3: 'I have never concealed from your Lordships that our mode of payment by results, as it is called, puts in the way of the good teaching and the good learning of these subjects almost insuperable obstacles' (Menet 1874: 13).

The next criticism concerned the increasingly mercenary approach to education engendered in school managers and teachers. By the end of the 1860s the term, 'farming of schools', was widely used in inspectors' and other official reports. Children were seen not as individuals or as learners, but as grant-earners. In some schools the teacher's salary was based directly upon the number of passes and grant earned. Indeed, in some cases, Menet reported, 'the commercial element is still further strengthened by the return of a certain amount in some shape to the children who pass in the Standards' (Menet 1874: 17).

Standards also interfered with the organisation of a school. In Menet's view, schools should be organised by classes, not by standards. Children should be grouped according to their educational needs, not according to grant-earning requirements. He also asked why standards and payment by results should be applied only to public elementary schools, but not to middle-class, private, grammar or public schools? What would a parent say to a master or mistress in one of these schools who announced that children would be limited to a narrow curriculum, not allowed to progress too quickly through that curriculum, and that only selected scholars who were likely to pass would be entered for the annual examination? Finally, Menet criticised 'the disastrous effects of the Standards . . . upon Inspection properly so called, as distinguished from mere Examination' (Menet 1874: 19).

Menet's solutions were based upon the abolition of the system of standards, and its replacement by a return to an annual examination by properly trained inspectors, 'an examination which would be fairly within the range of each school, according to its circumstances and standing' (Menet 1874: 23). He concluded:

> We want, on the one hand, less routine, less mechanism, less complication, fewer pains and penalties. We want in their place, on the other hand, a much fuller and clearer recognition of what Education really is, more freedom for Inspectors, more liberty for Teachers, more cultivation of mind, and more common sense. Let the weights be removed which press on all sides, and everybody concerned will breathe more freely. (Menet 1874: 26)

The criticisms of Menet and Payne, among others, did not go entirely unheeded. The system which they deplored was further modified, for example, by the addition of a seventh standard in 1882. In the same year, the Education Department's instructions to inspectors allowed them to calculate part of the grant upon 'the estimate you form of the merit of the school as a whole' (CCE 1883: 157). Although the quantity and quality of the pupils' passes were still to be the major factors in such an estimate, some allowance could also be made for 'special circumstances', for example 'a shifting, scattered, very poor or ignorant population' (CCE 1883: 157).

Payment by results came to an end in the last decade of the nineteenth century, but the concept and terminology of standards with which it had been associated, lasted into the twentieth. The very standards themselves, in tabular form, continued to be included in the elementary school code. In 1912 John Adams, first principal of the London Day Training College, delivered a presidential address entitled 'An objective standard in education' to the Educational Science section at the Dundee meeting of the British Association. In this address, and in his major work, *The Evolution of Educational Theory*, first published in the same year, Adams showed how the old concept of standards had so passed into common usage that people talked about children not as being of a certain age, or ability, but as 'being in standard so-and-so' (Adams 1912: 304). Adams gave a cautious welcome to the work of Binet and Simon in respect of intelligence scales, for, as he observed, standards, which had been 'Primarily meant as means of measuring the money value of the communication of certain bits of information . . . came in the teacher's hands to be a means of estimating ability' (Adams 1912: 304).

Conclusions

Four broad conclusions may be drawn.

The first is that the term, 'standards', has occupied a prominent and contested place in recent British educational history. This is not surprising, given the ever-changing nature and amount of knowledge, coupled with changes in educational and broader societal priorities. It is not difficult to dip into that history to find evidence for, or against, a decline in standards over time. On the one hand it is clear that literacy standards are higher in the twentieth century than in any previous period. Similarly, a much greater percentage of the population now

attends university and is educated to degree level than ever before. In the middle years of the twentieth century grammar schooling was deemed to be appropriate for twenty per cent of the population. Today thirty per cent enter higher education. Improved educational access and standards have been particularly visible in respect of females. Not until the nineteenth century did female literacy begin to equal that of males, nor women gain access to higher education. On the other hand, there can be little doubt that levels of attainment in some subjects were higher in the later nineteenth century than they are today. For example, it seems likely that: more children knew the Lord's Prayer, Creed and Catechism than today; more boys could work complex sums in multiplication and long division; more girls were skilled in needlework; more children could recite substantial amounts of poetry. Some of these accomplishments were the direct product of the system of standards. Under the Code of 1883, when English was taken as a class subject children in Standard I were required to 'repeat 20 lines of simple verse'. This was followed by '40 lines of poetry' with associated tasks for Standard II, 60 for III, 80 for IV, and 100 for V. Standard VI children had 'To recite 150 lines from Shakespeare or Milton, or some other standard author, and to explain the words and allusions' (CCE 1883: 132–3). It is not difficult to gather such evidence about rising or falling standards. What is more difficult, but considerably more worthwhile, is to situate and evaluate it in the context of the time, and in historical perspective.

The standards debate of the second half of the nineteenth century took place in a series of contexts which in some respects were quite different from those of today. But although there are differences, similarities are also apparent. The Revised Code was introduced in 1862, not in implementation of the major recommendation of the Newcastle Report, which called for the introduction of a local system of county boards and local inspectors, but to avoid it. Not until 1870 would local boards be allowed to intrude into the field of public education. In 1862 teachers, who in the eyes of the government had been getting above themselves, would immediately be brought to account by the introduction of 'a little free trade'. Payment by results, it was argued, meant that, in future, government and taxpayers would get better value for money. Elementary education would either be efficient or it would be cheap. But was elementary education made more efficient? Menet's most persuasive argument against the system of standards of his day was that many of the HMIs who operated the

assessment regime upon which payment by results was based, were prepared to state publicly that, on balance, it was harmful to good teaching and learning. They, and Menet himself, while fully committed to the need for some form of externally-based assessment, concluded that such assessment must be professional and formative as well as managerial and summative.

Evidence from the other side of the world lends support to this conclusion. For in New Zealand from 1877 an almost identical system applied. Its central features—a national syllabus divided into prescribed standards, school inspectors whose role was to apply the assessment system and take no responsibility for its effects—were essentially the same, although without the direct operation of payment by results. In New Zealand, school examination results were published in tabulated form in newspapers, and in 1880 the government declared that 'the school with the lowest average age and the highest percentage of passes in the same standards is the most efficient . . . . a high average age and a low percentage of passes indicates a school of the opposite character' (McKenzie 1994: 249). David McKenzie has recently concluded that the system of external review based upon standards as operated in New Zealand, even without the dimension of payment by results, was essentially miseducative. The initiatives of better teachers were curbed; weaker teachers taught to the test and some engaged in outright dishonesty. Intended minima soon became maxima; good teaching and educational improvement were stifled. McKenzie supports Matthew Arnold's identification of the underlying problem, namely that the fault lay in the bureaucratic system of evaluation itself, rather than in its specific use for the purpose of payment by results (McKenzie 1994: 251). McKenzie drew upon this historical perspective to argue that criticisms of the Education Review Office in New Zealand 'owe their origin to the failure of the Picot Committee to grasp that a distanced review authority which is required to act judgmentally will be unable to facilitate the co-operative activity that the process of educational review requires' (McKenzie 1994: 247).

The third point concerns the contributions which historians can make in relation to educational research. In 1996, in a lecture entitled 'Teaching as a research-based profession: possibilities and prospects', David Hargreaves argued that much educational research, unlike that in medicine or the natural sciences, was 'non-cumulative, in part because few researchers seek to create a body of knowledge which is then tested, extended or replaced in some systematic way' (Hargreaves 1996: 2). The

historian can employ a chronological approach to provide a cumulative account, but must also indicate some of the difficulties inherent in creating, testing, extending or replacing bodies of knowledge in educational research. There are continuities in educational history, and it is not difficult to show, for example, that some basic principles of teaching and learning have as much validity in the twentieth century as in the nineteenth. Progress in some matters—for example, better physical health and improved literacy rates—can also be demonstrated over long periods of time. But it is difficult to be prescriptive about all issues of educational practice, and to measure progress (or decline) in all areas, including overall educational standards, essentially because the concept of education (like that of progress) is not value-free. Education, indeed, has been well defined as initiation into worthwhile activities, and it is clear that assessments of what knowledge (and of what standards) are of most worth, have been, and will continue to be, matters of debate. It seems probable that one of the best means of ensuring improvements in educational standards is for that debate to be conducted in a constructive and co-operative way.

Finally, it is important to end on a positive note. The current government's commitment to the raising of educational standards is to be welcomed, as are many of its initiatives to achieve this aim, particularly the establishment of General Teaching Councils for England and Wales. Central government should continue to do what it can to contribute towards the raising of standards. But it should also recognize both the limitations of its own role, and that pupils, parents, teachers, local authorities and others have most significant parts to play.

Education is not susceptible to quick fixes, whether as a result of political intervention or pedagogical fashion. Teaching is not like some other professions, for example medicine and the law, where high profile and dramatic results may be easily and quickly achieved. The two fundamental factors in raising educational standards are the steady commitment to worthwhile education amongst pupils, parents and society at large, and the recruitment and retention of as many good teachers as possible.

This paper ends, as it began, with definitions. Worthwhile education is about the promotion of knowledge over ignorance, of truth over falsehood, of concern for others over selfishness, over effort over sloth, of mental and physical well-being over despair and debility. If we neglect these truths in order to put a spin or gloss upon education for other purposes—whether we are politicians, journalists, authors,

academics, teachers—then we shall be agents in lowering rather than raising educational standards. Good teachers may be defined as those with a sound knowledge of their subjects and of pedagogy, steady application of principles of management and organization, genuine concern for those whom they teach, and the ability to inspire and enthuse.

Discussion

Gillian Sutherland

Definitions

Richard Aldrich has rightly drawn attention to the range and elaboration of definitions of standards in the *Oxford English Dictionary.* The most relevant for an analysis of the appeals to standards in England over the last two centuries are those under the general heading II: 'Exemplar, measure of weight'. Sub-set 12 reads: 'A definite level of excellence, attainment, wealth or the like, or a definite degree of any quality, viewed as a prescribed object of endeavour or as the measure of what is adequate for some purpose.' The words 'prescribed' and 'adequate for some purpose' are deserving of emphasis: reminders that the aim must always be to reach behind the immediate appeal to standards and ask, 'for what purpose?' This must be an over-arching objective and a recurring theme in any extended discussion.

With this proviso, these dictionary definitions offer two linked but distinct notions. The first is that of a target of endeavour, a level of excellence, which only a few will reach. The second is that of a measure of what is adequate, a minimum acceptable level which almost everyone is able to reach. Between excellence and adequacy there may be a large gap. These two notions are nevertheless of considerable help in disentangling the various forms that invocation of standards took in the educational discourse of nineteenth century England. Such invocations were always linked to formal examinations, a mechanism embraced with passion by nineteenth-century educational activists; but formal examinations were conducted for a variety of objectives. The first wave of enthusiasm for examinations was part of a project to identify

and reward elites: standards were a measure of excellence. Only later, as schemes for mass education gained momentum, were examinations used and standards invoked to measure adequacy.

Nineteenth-century Discussions

(1) *Excellence and Elites*

The use of formal examinations to set a target of excellence which only a few would attain but representing a pinnacle against which others could measure themselves, gathered momentum in the universities of Oxford and Cambridge from the 1780s onward. By the end of the 1820s it was well-established, a model for other universities and being brandished as the tool to overhaul selection for the Indian Civil Service. Macaulay spoke eloquently in favour of using competitive examination to this end in the House of Commons in 1833 (Macaulay 1898: xi, 571–3). He justified his position in these terms:

> Education would be mere useless torture, if, at two or three and twenty, a man who had neglected his studies were exactly on a par with a man who had applied himself to them, exactly as likely to perform all the offices of public life with credit to himself and with advantage to society. Whether the English system of education be good or bad is not now the question. Perhaps I may think that too much time is given to the ancient languages and to the abstract sciences. But what then? Whatever be the languages, whatever be the sciences, which it is, in any age or country, the fashion to teach, the persons who will become the greatest proficients in those languages and those sciences will generally be the flower of the youth, the most acute, the most industrious, the most ambitious of honourable distinctions.

It took twenty years to replace patronage by competitive examination. However the experience was one on which Macaulay's brother-in-law, Charles Trevelyan, was able to draw when he was then asked to turn his attention to the Home Civil Service (Sutherland 1984: 97–100). By this time too examinations were being deployed to stir up moribund grammar school foundations and to assess the work of the new proprietary boarding schools. There was a bandwagon rolling on which women as well as men would climb in the second half of the century (Roach 1971, Fletcher 1980).

When the notion of a standard was invoked in such discussions it was almost always seen as an absolute, an external fixed reference point. In the University of Oxford a committee of the Hebdomadal Board, appointed in March 1829 to construct a new Examination Statute, saw

their objective to be the creation of a self-activating mechanism to stimulate both teaching and study and reaffirmed that 'the standard for each class be absolute and positive' (Brock and Curthoys 1998: 344). Theoretically all the candidates could be in the first class and individual classes could be empty; as, in subsequent years, they sometimes were. Those concerned with the technology of assessment in the late twentieth century would recognize an early example of criterion referencing. The criteria constructed in such examinations, moreover, trailed an aura of enormous power. Macaulay had made an explicit equation of ability and merit; and this equation was deployed enthusiastically and effectively by his supporters and successors (Sutherland 1984: 97–113, Young 1958).

(2) *Adequacy and Mass Schooling*

Developing from mid-century, side by side with but distinct from the invocation of excellence, was the invocation of a standard as a measure of adequacy, an acceptable minimum floor, which almost everyone was expected to achieve. Again examinations were the chosen mechanism and this is the discourse to which the Revised Code of 1862 and payment by results, explored at length by Aldrich, belonged.

Some of the immediate power and impact of the Revised Code and its examination Standards came from the fact that real money depended on them. For the thirty years of its operation the bulk of the government grant to a school—over half its annual income—depended on the children's passes in these Standards; and many managers tightened the link by making some fraction of the teacher's salary dependent on the pass rate achieved. Such acts dramatize the close connections between the Revised Code and the conceptual framework of free market economics.

(3) *The Market Model*

Robert Lowe, the Revised Code's principal architect, saw government as the consumer of mass education, getting—or failing to get—value for money in the form of certain minimum skill levels. He expounded his thinking in the House of Commons in February 1862 (*Hansard* 3rd ser. 13 February 1862: 205):

> What is the object of inspection? Is it simply to make things pleasant, give the schools as much as can be got out of the public purse, independent of their efficiency; or do you mean that our grants should not only be aids, subsidies and gifts, but fruitful of good? That is the question and it meets us at every

> turn. Are you for efficiency or for a subsidy? Is a school to be relieved because it is bad and therefore poor, or because it is a good school and therefore efficient and in good circumstances?

The use of the market as a model for educational provision was sharply challenged by one of Lowe's own inspectors, Matthew Arnold. After an initial challenge to the Revised Code itself, Arnold drew back: he could not afford to lose his job. He inspected, grumbling, under its rules for the remainder of his career (Sutherland 1973b: 12–13). However in writing about secondary education—his real passion and an area for which the state as yet took no responsibility—Arnold tackled the model of the market head on. In *A French Eton* in 1863 he proclaimed 'that to trust to the principle of supply and demand to do for us all that we want in providing education is to lean upon a broken reed' (Arnold 1863: 282). He continued:

> The mass of mankind know good butter from bad, and tainted meat from fresh, and the principle of supply and demand may, perhaps, be relied upon to give us sound meat and butter. But the mass of mankind do not so well know what distinguishes good teaching and training from bad; they do not here know what they ought to demand, and, therefore, the demand cannot be relied on to give us the right supply. Even if they knew what they ought to demand, they have no sufficient means of testing whether or no this is really supplied to them.

At this point Arnold was primarily concerned to challenge the idea that either pupils seeking education or their parents ought to be seen as consumers or customers. He went on, however, to argue that neither was the state a consumer or customer in the simple and straightforward sense that Lowe had posited. Rather, the state was pupils, parents, all citizens, in their collective and corporate character. Deliberately Arnold echoed Burke in seeing 'the citizens of a State, the members of a society' as a partnership, 'a partnership in all science, in all art, in every virtue, in all perfection'. Viewed from this standpoint, the provision of national education became for Arnold both more complex and a project of fundamental moral importance: one for which the model and language of the market were wholly inappropriate.

(4) *Standards and Averages*

Although the grant and inspection arrangements embodied in the Revised Code were dismantled in the 1890s, many of the habits and patterns formed by it persisted far longer. The teaching habits developed to survive within it died hard, as did the firm classification of

children by age and the physical structures developed to accommodate this grouping. How many readers of this note attended a primary school which labelled its classes—sometimes had incised above its doors—Standards I, II, III etc.? For the Revised Code had determined grant awards during the years of massive school building immediately following the 1870 Act.

Further, while 'Standard', remaining a class label, ceased to be an examination, the idea of a standard as a measure of adequacy was given fresh power and tied closely to the term 'average' by the work of Francis Galton. In his *Hereditary Genius* in 1869 and subsequently he argued that human abilities were distributed along a normal bell-shaped curve. Thus there would be a small number whose abilities and achievements were exceptional—excellent—and a small number whose abilities and achievements would be abysmal, while the abilities and achievements of the majority clustered around the middle ranks, the average (Mackenzie 1981: 56–8). Effective testing and examining could therefore be expected to spread candidates in the familiar bell-shaped pattern: a handful at each extreme, the majority bunched in the middle.

This assumption was fundamental to the subsequent development of group mental, or as they came to be called, intelligence tests and of standardized tests of attainments (Sutherland 1984: 115–27). Like the tests of the Revised Code, these are tests for ordering large populations and were used as such in secondary school selection in England from the 1920s onwards (Sutherland 1984: 164–269, Thom 1986: 117–23). Standards in this discourse were what technologists of assessment would call norm referenced, shaped by the pattern of performance of the majority. Yet at the same time the assumption that human abilities follow the normal curve of distribution made it easy to bring measures of excellence and measures of adequacy into a close linear relationship.

The Continuing Resonances of Past Debates

Richard Aldrich has noted some of the criticisms of payment by results made at the time. These could be extensively amplified from the reports of other inspectors (e.g. Sutherland 1973: 195–7) and from the comments of a later chief inspector, Edmond Holmes, on the revelation it was to him when payment by results came to an end. His lament that the methods of teaching and examining reading created a culture in which children left school never having learnt to use books (Holmes 1911: 127), has an extraordinarily contemporary ring. Worries

of the 1990s about the costs and administrative burden of an elaborate apparatus of national assessment evoke powerful echoes of the infighting within the Civil Service in the 1880s over precisely such issues—and this infighting was one factor contributing to the destruction of payment by results (Sutherland 1973: 237–45).

Exact analogies, however, are always difficult to sustain and comparisons forward are dangerously a-historical. An intellectual tactic which is more respectable and robust than such comparisons is to treat the nineteenth century's invocation of 'standards' as a piece of vicarious experience. This prompts several questions which seem pertinent for the late twentieth century debate. Are standards understood to be the target of excellence, which only the few can attain, although they may define and rank the lesser achievements of the many? Or are they meant to be the minimum acceptable floor for almost everyone? Or does popular discourse slither in confused fashion back and forth between these two poles? Has an average now become this minimum acceptable floor, something everyone must achieve, detached from any conceptual relation it once had to a normal curve of distribution? These seem appropriate questions to ask when a successor to Edmond Holmes heads the Office for Standards in Education; and when publication of A level results, GCSE results and primary school league tables generates regular media feeding frenzies.

Finally we should return to questions of purpose. What models are the appropriate ones to apply to the provision of education? This seems a pertinent question when the possibility of handing over schools, both 'failing' and functioning, to entrepreneurs begins to be canvassed (e.g. Davis 1993). Whenever the cry of 'standards' is invoked, the supporting agenda must be explored too.

Anthony Heath

I enjoyed Aldrich's paper and learned a great deal from it, particularly about the Revised Code which clearly has intriguing parallels with contemporary attempts to secure better value for money in education.

Aldrich makes a crucial distinction between the notion of a standard as a yardstick for judging performance and a standard in the sense of the average level of attainment as measured by that yardstick. While there is bound to be considerable uncertainty about the equivalence of the yardsticks used in different periods, there seems little doubt that

there has been historically a substantial improvement in the educational attainments of the British population, at least in the sense that larger proportions of the population have acquired basic literacy and formal qualifications. Of course in this respect Britain is no different from other industrialised countries (see for example Blossfeld and Shavit 1993).

Aldrich describes a steady improvement in literacy across five centuries in Britain, although this improvement has not perhaps been maintained in the second half of the twentieth century. However, lack of recent progress on raising standards, that is the lack of growth in the proportion of the age-group reaching a basic level of literacy, has perhaps been compensated for by increasing proportions reaching higher levels of educational attainment such as GCSE, A level and degree.

This can be illustrated with data from the General Household Surveys. The GHS collects data on respondents' highest educational qualifications, and we can arrange these data by the respondents' years of birth. Arranged in this way the GHS data give a picture of the growth over the course of the twentieth century in the proportions from successive birth cohorts who have acquired educational qualifications. (Since people can go on acquiring qualifications throughout the life cycle, these cohort differences almost certainly underestimate the true changes in the qualifications of the population at the time they left full-time education.)

As we can see, the percentage with no formal educational qualification fell from 73% of men and 86% of women in the oldest birth cohort, born in the years 1900–09, to 12% of men and 16% of women in the youngest cohort, born seventy years later. Perhaps the most striking increases are in the percentages with intermediate qualifications such as A or O level (and their historical equivalents). The increases at these levels have been substantially greater than those at the highest qualification levels such as degree.[1]

If we take basic literacy as the lowest level of attainment, as Aldrich notes, 'take-off' occurred in the nineteenth century or earlier, and began to approach a ceiling in the second half of the twentieth century. The next expansion, as shown in Table 1, was in intermediate secondary qualifications which 'took off' much later, in the second half of the twentieth century, and may be expected to reach their ceiling early in the twenty-first century. A third expansion, in tertiary education, has only recently begun and it is too early to say when and at what level a ceiling will be reached.

Table 1. Highest qualification by cohort for males and females (%)

	1900–09	1910–19	1920–29	1930–39	1940–49	1950–59	1960–69	1970–79
Males								
Degree	3.0	4.0	5.9	7.7	10.5	12.6	11.5	9.4
Higher: below degree	2.9	3.7	5.7	7.6	10.5	12.1	12.0	11.5
A level	1.4	2.1	3.0	5.8	9.7	16.4	18.3	30.9
O level	3.8	5.7	6.7	11.1	15.9	20.0	26.5	27.3
Low	16.2	17.9	18.3	18.4	14.0	10.6	14.1	8.6
None	72.8	66.6	60.3	49.3	39.4	28.4	17.6	12.2
N	3101	15900	29444	29334	35284	33945	21441	3949
Females								
Degree	.9	.9	1.4	2.4	4.0	6.9	8.7	11.4
Higher: Below degree	3.2	4.3	5.2	7.9	8.8	10.1	8.3	5.4
A level	.8	.7	1.0	2.0	3.8	7.7	12.1	16.8
O level	2.6	4.1	6.4	10.4	16.8	25.4	36.4	38.9
Low	6.4	9.1	11.7	13.7	16.6	14.5	16.4	11.4
None	86.2	80.9	74.4	63.6	50.0	35.5	18.1	16.1
N	3866	19440	33259	31655	38371	36881	23185	4156

Notes.

1. To maximize the reliability and time-span of the estimates we have cumulated General Household Surveys for the years 1973 to 1992. We restrict table 1 to respondents aged 21 and over at the time of the survey in which they were interviewed.
2. The decline in the youngest birth cohorts in the proportions obtaining a degree reflects the fact that degrees are typically obtained at older ages and that some members of these younger cohorts had not yet finished their education.

In describing the trends in attainment over the twentieth century, it must be recognised that we cannot make any claims to exact equivalence in the yardsticks used. While in some cases, such as degrees, there has been some institutional continuity at least in the title of the qualification, in others there have been major reorganisations. For example, school Certificate and Higher School Certificate were replaced by GCE O and A level; the new qualification of CSE was introduced and then amalgamated with O level to form the GCSE. There are conventions about what count as equivalent: for example, a pass at school certificate is taken to be equivalent to an O level at grades 1–5 or, subsequently, at grades A–C; CSE grade 1 and GCSE at grades A–C are assumed to be the more recent equivalents. All can be regarded as intermediate secondary qualifications taken at around age 15/16.

We certainly could not claim that this convention provides exact equivalence. It is quite possible to argue that in some sense a grade C at GCSE today is at a lower standard than a grade C at O level and that the yardstick has in this way been debased. Some of the growth in the percentages gaining O level equivalents, as shown in table 1, might thus be accounted for by the 'standards' becoming easier. On the other hand we would want to claim that this convention is better than most alternative conventions: it would make even less sense to treat *any* pass at O level and GCSE as equivalent irrespective of the grade obtained. In other words, the convention being used is probably the least bad of the possible conventions, but certainly falls short of the ideal (although the exact nature of the ideal may itself be unknowable). Moreover, I doubt if anyone would wish to argue that the whole of the increase in attainments shown in Table 1 could be explained by debasement of the yardstick. Indeed, if we adopted a more conservative convention and claimed that a contemporary A level is of the same standard as the pre-war school certificate, the GHS data would still show a huge increase in attainment.

Of course, failure to maintain comparable yardsticks over time is a serious handicap for sociologists who wish to determine whether social class, gender or ethnic inequalities in attainment have been reduced over time or who wish to monitor the effectiveness of government education policy (see for example Heath and Clifford 1990). If our yardstick changes, it becomes difficult to provide definitive answers to our research questions. But as Jencks once remarked 'We are aware of the hazards involved and have tried to check the validity of our assumptions wherever possible. Nonetheless, the methods we have used may involve considerable error. In self-defence, we can only say that the magnitude of these errors is almost certainly less than if we had simply consulted our prejudices, which seems to be the usual alternative' (Jencks 1972: 15).

But should we even expect the yardstick to remain the same over time? As Aldrich argues, the societal functions have changed over time and there may be good educational reasons for the changing nature of the yardsticks used. One function, although not perhaps the only one, is that these qualifications are used as a selection device by, for example, schools, universities and to a lesser extent employers.

However, as secondary education has expanded and more students stay on to take the public examinations at age sixteen, so the purpose of the selection has also changed. Thus pre-war, when very few pupils

stayed on at secondary school beyond the minimum leaving age (which was then fourteen), school certificate and matriculation were geared to University entrance. After the war, as increasing numbers stayed on at school, O levels became a prerequisite for access to A level courses in the sixth form. After the raising of the school-leaving age in 1973/4 O level, CSE and now GCSE had to cater for the whole ability range and began to function as a school-leaving certificate for many pupils. Given the huge changes in numbers staying on at school and taking these examinations, and given the very different 'gate-keeper' roles that these examinations performed, it would be extraordinary if they had maintained the same 'standards' over time. And it would probably have been quite inappropriate if they had done so: a yardstick that would have been a good discriminator between borderline candidates for university admission before the war is unlikely to be a good discriminator for borderline candidates for entry to the sixth form after the war or for entry to skilled occupations at the end of the century.

Given the historical transformation of our educational system, therefore, and the changing selective functions which the public examinations have been asked to perform, it is highly desirable that the yardstick should be changed from time to time so that it is appropriate for its current function. If what we want is a selection device, then we need something that is a good discriminator at the borderline. As the borderline changes, so should the yardstick.

Public examinations are a central component of a meritocratic selection procedure for selecting young people for post-school entry to colleges or employment. It follows that their primary purpose should be to secure equity between current applicants for entry rather than between generations. Rather than debating whether yardsticks have been debased over time, it is more useful to consider whether current examination procedures are adequately designed for the selection functions they are currently asked to perform.

Acknowledgements

I would like to thank the Office of National Statistics and the Data Archive at the University of Essex for making the General Household Survey data available. I am very grateful to Jane Roberts at the Social Studies Computing Unit, Oxford, for her work on cumulating the files.

Sig Prais

The specification of schooling standards on a *nationwide* basis is not something newly devised in the past decade for the National Curriculum in this country. Richard Aldrich has reminded us in his paper, that it can be traced back to the middle of the last century in relation to core subjects. I should like to offer some comments on what we mean, or should mean, when we talk of educational standards. The clarification of a number of complex associated issues, I shall suggest, would do much to improve rational policy formation. My comments will focus on the gap between specified standards and pupils' actual attainments, the meaning of expected standards and the care needed in measuring pupils at the lower end of the achievement spectrum.

First, the gap between specified *standards* and pupils' actual *attainments* is important. For example, does a teacher concentrate on bringing more marginal low-attaining children across a lower boundary (as under the English nineteenth century payment-by-results system), or does she concentrate on the highest attaining children, to bring more up to those higher standards on which her success—and that of her school—may be assessed, as in the days of the 11–plus and scholarship examinations? That difference in teaching emphasis is well-recognised.

Secondly, expectations: when government policy speaks of the educational 'standards of 11 year-olds *expected* for their age' we need to be very clear as to what is meant by 'expected'. Statisticians use the term 'expectation' to mean simply the arithmetic average, without any moral or policy implications: a statistician might say that the 'expected' height of a grown man in England is 5 foot 7 inches. There is no necessary implication that those below the average could or should be raised to that 'expectation' by a series of policies, even if that were possible. If educational policies now aim to raise 80 per cent of pupils, as is said, to the 'standards expected for their age', it is probably most reasonable to interpret this as meaning that at some future date the top 80 per cent of pupils will reach the attainments that the middle pupil happens to reach today.

To require 80 per cent of children to reach standards in literacy and numeracy hitherto attained only by 50 per cent seems likely to require a great shift in school time-tables towards those subjects (English and mathematics), and away from other subjects, together with reforms in syllabuses, teaching methods and classroom organisation. Some of those reforms are under way in the new Labour Government's policy

recommendations for primary schools for a daily Literacy Hour and a Numeracy Hour; but prudent observers will hold their breath as to whether present and planned reforms in teaching methods and classroom organisation will go far enough to achieve the stated aims.

In moving the focus of official educational policy towards pupils at the lower part of the attainment range, we may detect a belated recognition of the view that paramount importance in current schooling reforms attaches to the *employment* consequences of continuing advances of automation and computerisation throughout the economy. The increased demand for personnel to serve as highly skilled technicians has long been clear; but of equal economic significance, and probably of greater social significance, is the decline in employability of that great proportion of school-leavers who previously were provided with schooling which led only to unskilled work.

An associated worry arises from the spread of new styles of teaching which incorporate a greater degree of 'discovery learning' by the pupil— such as, learning to read by recognising the length and shape of whole words rather than how sounds are represented by letters ('look-and-say' versus 'phonics'), or relying on a calculator at early ages rather than embedding arithmetical bonds in the mind (Bierhoff 1996: 152). There is space here only to adumbrate the complexities at issue: it is possible that these new methods may work well with high-ability children (the methods were often developed in 'model schools' attached to universities, attended by professors' children with all the help that such children have at home at their disposal); but they may have done a great disservice to children from problem homes who rely to a greater extent on their teachers for guidance, and require a stable classroom environment for emotional security as a pre-condition for efficient learning. The new teaching methods may thus have done both some good and some harm; irrespective of whether the *average* has slightly risen or slightly fallen, we need to have our eye on the *spread* of attainments a matter to which educational researchers have given little attention. The great worry is that, in an era when technological developments in the economy have reduced employment opportunities for children of below-average attainments, developments in teaching may have served to reduce their opportunities yet further.

Bibliography

Adams, J. (1912). *The Evolution of Educational Theory* (London, Macmillan).

AIE (1996). *Assessment in Education*, **3**(2).

Aldrich, R. (1995). *School and Society in Victorian Britain: Joseph Payne and the new world of education* (New York, Garland).

Aldrich, R. (1996). *Education for the Nation* (London, Cassell).

Aldrich, R. (1997). *The End of History and the Beginning of Education* (London, Institute of Education).

Aldrich, V. C. (1963). *Philosophy of Art* (Englewood Cliffs, Prentice-Hall).

Anderson, R. D. (1995). *Education and the Scottish People* (Oxford, Oxford University Press).

Arnold, Matthew (1863). *A French Eton*, reprinted in *The Complete Prose Works of Matthew Arnold vol. ii, Democratic Education* ed. R. H. Super (Ann Arbor, University of Michigan Press, 1962), pp 262–325.

Arnott, M. (1993). Thatcherism in Scotland: an Exploration of Educational Policy in the Secondary Sector (PhD Thesis, Strathclyde University).

Ayer, A. J. (1946). *Language Truth and Logic* Second edition (London; Penguin).

Baird, J. (1998). What's in a Name? Experiments with blind marking in A-level Examinations. *Educational Research*, **40**(2), 191–202.

Baird, J. and Jones, B. (1998). Statistical analyses of examination standards: better measures of the unquantifiable? (Associated Examining Board Research Report—RAC/780).

Bardell, G.; Fearnley, A. and Fowles, D. (1984). *The contribution of graded objectives schemes in Mathematics and French* (Manchester, Joint Matriculation Board).

Barnes, B. (1974). *Scientific Knowledge and Sociological Theory* (London, Routledge and Kegan Paul).

Bartholomew, D. J. and Knott, M. (1999). *Latent Variable Models and Factor Analysis* (2nd edition) (London, Arnold).

Bartholomew, D. J. (1996). *The Statistical Approach to Social Measurement* (San Diego, Academic Press).

Beardsley, M. C. (1981). *Aesthetics: Problems in the Philosophy of Criticism* (Indianapolis, Hackett).

Beaton, A. E. and Zwik, R. (1990). *Disentangling the NAEP 1985–86 reading anomaly.* (Princeton, Educational Testing Service).

Benn, C. and Chitty, C. (1996). *Thirty Years On* (London, David Fulton).

Berger, P. and Luckmann, T. (1966). *The Social Construction of Reality* (London, Penguin).

Berry, C. (1997). *Social Theory of the Scottish Enlightenment* (Edinburgh, Edinburgh University Press).

Best, D. (1985). *Feeling and Reason in the Arts* (London, Allen & Unwin).

Bierhoff, H. (1996). Laying the foundation of numeracy: a comparison of primary

school textbooks in Britain, Germany and Switzerland. *Teaching Mathematics and Its Applications*, **15**, 141–60.

Billington, R. (1988). *Living Philosophy: An Introduction to Moral Thought* (London, Routledge).

Bourdieu, P. (1989). *La Noblesse d'État: Grandes Écoles et Esprit de Corps* (Paris, Les Editions de Minuit).

Brock, M.G. and Curthoys, M.C. (1998). (eds.). *The History of the University of Oxford vol. vi, Nineteenth-Century Oxford, Part 1* (Oxford, Clarendon Press).

Brooks, G. (1997). Trends in standards of literacy in the United Kingdom, 1948–1996 (paper presented at the UK Reading Association conference, University of Manchester, July 1997, and at the British Educational Research Association conference, University of York, September 1997).

Brown, A., McCrone, D., Paterson, L. and Surridge, P. (1998). *The Scottish Electorate* (London, Macmillan).

Burnhill, P., Garner, C. and McPherson, A. (1990). Parental education, social class and entry to higher education, 1976–1986. *Journal of the Royal Statistical Society*, series A, **153**, 233–248.

Burstein, J., Kaplan, R., Wolff, S., and Chi, L. (1997). Using Lexical Semantic Techniques to Classify Free-Responses (Princeton N.J. Educational Testing Service Research Report available on ETSnet at http: //www.ets.org/research/siglex.html).

Christie, T. and Forrest, G. M. (1981). *Defining Public Examination Standards* (London, Schools Council/Macmillan).

Cipolla, C. M. (1969). *Literacy and Development in the West* (London, Penguin).

Clanchy, M. (1979). *From Memory to Written Record: England 1066–1307* (London, Edward Arnold).

Collins, R. (1979). *The Credential Society* (New York, Academic Press).

Committee of Council on Education (1863). *Report of the Committee of Council on Education 1862–63* (London).

Committee of Council on Education (1872). *Report of the Committee of Council on Education 1871–72* (London).

Committee of Council on Education (1873). *Report of the Committee of Council on Education 1872–73* (London).

Committee of Council on Education (1883). *Report of the Committee of Council on Education 1882–83* (London).

Copeland, J. (1993). *Artificial Intelligence: A Philosophical Introduction* (Oxford, Blackwell).

Cox, C. B. and Dyson, A. E. (1971). (eds.). *The Black Papers on Education* (London, Davis-Poynter).

Cresswell, M. J. (1987). Describing Examination Performance: grade criteria in public examinations. *Educational Studies*, **13**(3), 247–65.

Cresswell, M. J. (1990). Gender Effects in GCSE—Some Initial Analyses (Paper prepared for a Nuffield Seminar at University of London Institute of Education on 29 June 1990) (Unpublished Associated Examining Board Research Report—RAC/517).

Cresswell, M. J. (1994). Aggregation and Awarding methods for National Curriculum

Assessments in England and Wales: a comparison of approaches proposed for Key Stages 3 and 4. *Assessment in Education*, **1**(1), 45–61.

Cresswell, M. J. (1995). Technical and Educational Implications of using Public Examinations for Selection to Higher Education. In T. Kellaghan (ed.), *Admission to Higher Education: Issues and Practice* (Dublin, Educational Research Centre and Princeton, International Association for Educational Assessment).

Cresswell, M. J. (1996). Defining, Setting and Maintaining Standards in Curriculum Embedded Examinations: Judgemental and Statistical Approaches. In H. Goldstein and T. Lewis (eds.). *Assessment: Problems, Developments and Statistical Issues* (London, Wiley).

Cresswell, M. J. (1997a). *Examining Judgements: Theory and Practice of Awarding Public Examination Grades* (PhD thesis, University of London Institute of Education).

Cresswell, M. J. (1997b). Can Examination Grade Awarding be Objective and Fair at the Same Time? Another Shot at the Notion of Objective Standards (Unpublished Associated Examining Board Research Report—RAC/733).

Cresswell, M. J. and Houston, J. G. (1991). Assessment of the National Curriculum—some fundamental considerations. *Educational Review.* **43**, 63–78.

Cressy, D. (1980). *Literacy and the Social Order: reading and writing in Tudor and Stuart England* (Cambridge, Cambridge University Press).

Damasio, A. R. (1995). *Descartes Error: Emotion, Reason and the Human Brain* (London, Papermac).

Davis, E. (1993). *Schools and the State* (London, Social Market Foundation).

Dean, C. (1998). Standards are not parents' top priority. *Times Educational Supplement*, 9 October.

Dearing, R. (1995). *Review of the 16–19 qualifications* (London, Department of Education).

Dennett, D. (1993). *Consciousness Explained* (London, Penguin).

Department for Education and Employment (DfEE). (1997). *Excellence in Schools* (London, Stationery Office).

Department of Education and Science (1967). *Children and Their Primary Schools. A Report of the Central Advisory Council for Education (England).* ii (London, DES).

Devine, M., Hall, J., Mapp, J. and Musselbrook, K. (1996). *Maintaining Standards: Performance at Higher Grade in Biology, English, Geography and Mathematics* (Edinburgh, Scottish Council for Research in Education).

Devlin, K. (1997). *Goodbye Descartes: The End of Logic and the Search for a New Cosmology of the Mind* (New York, Wiley).

Dore, R. (1996). *The Diploma Disease.* 2nd edition (London, Institute of Education).

Dreyfus, H. L. (1992). *What computers still can't do: a critique of artificial reason.* (Cambridge Mass., MIT Press).

Eagleton, T. (1993). *Literary Theory: An Introduction* (Oxford, Blackwell).

Eiser, J. R. (1990). *Social Judgement* (Milton Keynes, Open University Press).

Elwood, J. and Comber, C. (1996). *Gender differences in examinations at 18+* (London, Institute of Education).

Firestone, W. A. (1998). A Tale of Two Tests: Tensions in Assessment Policy. *Assessment in Education*, **5**(2), 175–192.

Fletcher, S. (1980). *Feminists and Bureaucrats. A study in the development of girls' education in the nineteenth century* (Cambridge, Cambridge University Press).

Fogelin, R. J. (1967). *Evidence and Meaning: Studies in Analytic Philosophy* (London, Routledge).

Forrest, G. M. and Orr, L (1984). *Grade Characteristics in English and Physics* (Manchester;,Joint Matriculation Board).

Foxman, D., Ruddock, G. and McCallum, I. (1990). *APU mathematics monitoring 1984–88 (Phase 2)* (London, Schools Examination and Assessment Council).

Fremer, J. (1989). Testing Companies, Trends and Policy Issues: A current view from the testing industry. In B. R. Gifford (ed.), *Test Policy and the Politics of Opportunity Allocation: The Workplace and the Law* (Boston, Kluwer).

French, S., Slater, J. B., Vassiloglou, M. and Willmott, A. S. (1987). *Descriptive and Normative Techniques in Examination Assessment* (Oxford, UODLE).

Galton, M. (1998). Back to consulting the ORACLE. *Times Educational Supplement*, 3 July.

Gierl, M. J. and Rogers, W. J. (1996). Factor analysis of the Test Anxiety Inventory using Canadian high school students. *Educational and Psychological Measurement*, **56**, 315–324.

Goldstein, H. (1983). Measuring Changes in Educational Attainment Over Time: Problems and Possibilities. *Journal of Educational Measurement*, **20**, 369–78.

Goldstein, H. (1995). *Interpreting International Comparisons of Student Achievement* (Paris, UNESCO).

Goldstein, H. (1996a) (ed.). *Assessment in Education, 3,* 2. Special Issue: The IEA Studies.

Goldstein, H. (1996b). International Comparisons of Student Achievement. In Little and Wolf (1996).

Goldstein, H. (1999). Performance Indicators in Education. In D. Dorling and S. Simpson (eds.). *Statistics in Society* (London, Arnold).

Goldstein, H. and Cresswell, M. J. (1996). The comparability of different subjects in public examinations: a theoretical and practical critique. *Oxford Review of Education*, **22**(4), 435–42.

Goldstein, H. and Wood, R. (1989). Five decades of item response modelling. *British Journal of Mathematical and Statistical Psychology*, **42**, 139–167.

Good, F. J. and Cresswell, M. J. (1988a). *Grading the GCSE* (London, Secondary Examinations Council).

Good, F. J. and Cresswell, M. J. (1988b). *Differentiated Assessment: Grading and Related Issues* (London, Secondary Examinations Council).

Gould, S.J. (1984). *The Mismeasure of Man* (London, Penguin).

Gray, J., Hopkins, D., Reynolds, D., Wilcox, B., Farrell, S. and Jesson, D. (1999). *Improving Schools: Performance and Potential* (Milton Keynes, Open University Press).

Gray, J., McPherson, A. and Raffe, D. (1983). *Reconstructions of Secondary Education* (London, Routledge).

Green, A., Leney, T. and Wolf, A. (1997). *Convergences and Divergences in European Education and Training Systems* (Brussels, EC Directorate-General XXII (Education, Training and Youth)).

Green, A., Wolf, A. and Leney, T. (1999). *Convergence and Divergence in European Education and Training Systems* (London, Institute of Education).

Hacking, I. (1965). *The Logic of Statistical Inference* (Cambridge, Cambridge University Press).

Hacking, I. (1990). *The Taming of Chance* (Cambridge, Cambridge University Press).

Hambleton, R. K. and Zaal, J. N. (eds.) (1991). *Advances in Educational and Psychological Testing* (Boston, Kluwer).

Hargreaves, D. H. (1996). Teaching as a research-based profession: policies and prospects (Teacher Training Agency annual lecture).

Heath, A. F. and Clifford, P. (1990). Class inequalities in education in the twentieth century. *Journal of the Royal Statistical Society*, series A, **153**, 1–16. .

Holland, P. W. and Rubin, D. B. (1982). *Test Equating* (New York, Academic Press).

Hollis, M. and Lukes, S. (1982). (eds). *Rationality and Relativism* (Oxford, Blackwell).

Holmes, E. (1911). *What Is and What Might Be* (London, Constable).

Jencks, C. (1972). *Inequality: A Reassessment of the Effect of Family and Schooling in America* (New York, Basic Books).

Johnson, V. E. (1997). An alternative to the traditional GPA for evaluating student performances. *Statistical Science*, **12**, 251–278.

Kelly, A. (1976). A study of the comparability of external examinations in different subjects. *Research in Education*, **16**, 37–63.

Kilpatrick, J. and Johansson, B. (1994). Standardised Mathematics Testing in Sweden: The Legacy of Frits Wigforss. *Nordic Studies in Mathematics Education*, **1**, *6–30.*

Koretz, D., Broadfoot, P. and Wolf, A. (1998) (eds.). *Assessment in Education*, **5**(3) (Special Issue on Portfolios and Records of Achievement).

Kuhn, T. (1970). *The Structure of Scientific Revolutions*, second edition (Chicago, University of Chicago Press).

Lakatos, I. (1974). *Proofs and Refutations: the Logic of Mathematical Discovery* (Cambridge, Cambridge University Press).

Little, A. (1996) (ed.). *Assessment in Education*, *4*(1) (Special Issue: The Diploma Disease Twenty Years On).

Little, A., Wang Gang, and Wolf, A. (1995) (eds.). *Sino-British Perspectives on Educational Assessment* (London, ICRA, Institute of Education).

Little, A. and Wolf, A. (1996) (eds.). *Assessment in Transition: Learning, monitoring and selection in international perspective* (Oxford, Pergamon).

Long, H. A. (1985). Experience of the Scottish Examinations Board in developing a grade-related criteria system of awards (Paper presented at the 11th annual conference of the International Association for Educational Assessment held in Oxford, England).

Macaulay, Lord (1898). *Collected Works*, 12 vols. (London, Longmans Green).

Mackenzie, D. A. (1981). *Statistics in Britain 1865–1930. The Social Construction of Scientific Knowledge* (Edinburgh, Edinburgh University Press).

McKenzie, D. (1994). The irony of educational review. *New Zealand Annual Review of Education*, **4**, 247–59.

McLean, L. D. (1996). Large-Scale Assessment Programmes in Different Countries

and International Comparisons. In H. Goldstein and T. Lewis (eds.), *Assessment: Problems, Developments and Statistical Issues* (Chichester, Wiley).

McPherson, A. and Willms, J. D. (1987). Equalisation and improvement: some effects of comprehensive reorganisation in Scotland. *Sociology*, **21**, 509–39.

Madaus, G. and Raczek, A. (1996). Turning Point for Assessment: Reform Movements in the United States. In Little and Wolf (1996).

Menet, J. (1874). *A Letter to a Friend on the Standards of the New Code of the Education Department* (London, Rivingtons).

Morrison, H. G., Busch, J. C. and D'arcy, J. (1994). Setting reliable national curriculum standards: a guide to the Angoff procedure. *Assessment in Education*, **1**, 181–199.

Murphy, R. J. L. (1982). Sex differences in Objective Test performance. *British Journal of Educational Psychology*, **52**, 213–19.

Murphy, R. J. L., Burke, P., Cotton, T., Hancock, J., Partington, J., Robinson, C., Tolley, H., Wilmut, J. and Gower, R. (1996). *The Dynamics of GCSE Awarding: Report of a project conducted for the School Curriculum and Assessment Authority* (London, SCAA).

Newcastle Report (1861). *Report of the Commissioners appointed to inquire into the State of Popular Education in England*, PP 1861 XXI (ii) (London).

Newton, P. (1996). The reliability of marking of GCSE scripts: Mathematics and English. *British Educational Research Journal*, **22**, 405–20.

Newton, P. (1997a). Measuring comparability of standards between subjects: why our statistical techniques do not make the grade. *British Educational Research Journal*, **23**(4), 433–49.

Newton, P. (1997b). Examining Standards Over Time *Research Papers in Education*, **12**(3), 227–48.

Orr, L. and Forrest, G. M. (1984). *Investigation into the relationship between grades and assessment objectives in History and English examinations* (Manchester, Joint Matriculation Board).

Orr, L. and Nuttall, D. L. (1983). *Determining Standards in the Proposed Single System of Examinations at 16+* (London, Schools Council).

Paterson, L. (1992). The influence of opportunity on aspirations among prospective university entrants from Scottish schools, 1970–1988. *Statistics in Society, Journal of the Royal Statistical Society*, series A, **155**, 37–60.

Paterson, L. (1995). Social origins of under-achievement among school-leavers. In L. Dawtrey, J. Holland, M. Hammer and S. Sheldon (eds.), *Equality and Inequality in Education Policy* (Milton Keynes, Open University Press).

Paterson, L. (1997). Student achievement and educational change in Scotland, 1980–1995. *Scottish Educational Review*, **29**, 10–19.

Paterson, L. (1998). The Scottish parliament and Scottish civil society: which side will education be on? *Political Quarterly*, **69**, 224–33.

Paterson, L. (forthcoming). Scottish traditions in education. In H. Holmes (ed.), *Compendium of Scottish Ethnology, vol. 11* (Edinburgh, Scottish Ethnological Research Centre).

Paterson, L. and Raffe, D. (1995). Staying on in full-time education in Scotland. *Oxford Review of Education*, **21**, 3–23.

Payne, J. (1872). 'Why are the Results of our Primary Instruction so Unsatisfac-

tory?', *Transactions of the National Association for the Promotion of Social Science.*

Phillips, M. (1996). *All Must Have Prizes* (London, Little, Brown and Company).

Pirsig, R. M. (1974). *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values* (London, Bodley Head).

Plewis, I. (1998). Inequalities, Targets and Zones. *New Economy*, **5**, 104–8.

Plewis, I. (1999). What's Worth Comparing in Education? In D. Dorling and S. Simpson (eds.). *Statistics in Society* (London, Arnold), 273–80.

Pole, D. (1961). *Conditions of Rational Inquiry: A Study in the Philosophy of Value* (London, Athlone).

Power, M. (1997). *The Audit Society: Rituals of Verification* (Oxford, Oxford University Press).

QCA (1998). *GCSE and GCE A/AS code of practice* (London, Qualifications and Curriculum Authority).

Reynolds, D., Creemers, B. P. M., Stringfield, S. and Teddlie, C. (1998). Climbing an educational mountain: conducting the International School Effectiveness Research Project. In G. Walford, *Doing research about education* (Lewes, Falmer Press).

Roach, J. P. C. (1971). *Public Examinations in England 1850–1900* (Cambridge, Cambridge University Press).

Robertson, C. (1992). Routes to higher education in Scotland. *Scottish Educational Review*, **24**: 3–16.

Rose, S. (1997). *Lifelines, Biology, Freedom, Determinism* (London, Penguin).

Sadler, D. R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education*, **13**, 191–209.

Schools Council (1979). *Standards in Public Examinations: Problems and Possibilities*, Report from the Schools Council Forum on Comparability (London, Schools Council).

SEC (1984). *The development of Grade-related Criteria for the General Certificate of Secondary Education—a briefing paper for working parties* (London, Secondary Examinations Council).

SEC (1985). *Reports of the Grade-related Criteria Working Parties* (London, Secondary Examinations Council).

SEC (1986). Draft Grade Criteria. *SEC News Number 2* (London, Secondary Examinations Council).

SEC (1987). Grade Criteria—Progress Report. *SEC News Number 6* (London, Secondary Examinations Council).

Shavit, Y. and Blossfeld, H. P. (1993). *Persistent Inequality: Changing Educational Attainment in Thirteen Countries* (Boulder, Col., Westview Press).

Skolöverstyrelsen (1980). Quoted in J. Kilpatrick and B. Johansson (1994). Standardised Mathematics Testing in Sweden: The legacy of Frits Wigforss. *Nordic Studies in Mathematics Education*, **1**, 6–30. .

Smith, J. V. and Hamilton, D. (1980) (eds). *The Meritocratic Intellect* (Aberdeen, Aberdeen University Press).

Start, B. and Wells, K. (1972). *The trend of reading standards* (Slough, National Foundation for Educational Research).

Stedman, L. C. (1998). An Assessment of the Contemporary Debate over US

Achievement. In D. Ravitch (ed.), *Brookings Papers on Education Policy* (Washington DC, Brookings Institution Press), 53–119.

Stephens, W. B. (1987). *Education, Literacy and Society, 1830–70: the geography of diversity in provincial England* (Manchester, Manchester University Press).

Sutherland, G. (1973a). *Policy-Making in Elementary Education 1870–1895* (Oxford, Clarendon Press).

Sutherland, G. (1973b) (ed.). *Matthew Arnold on Education* (London, Penguin).

Sutherland, G. (1984). *Ability, Merit and Measurement. Mental testing and English education 1880–1940* (Oxford, Clarendon Press).

The Scotsman Education (1998). 30 September: 4–5.

Thom, D. (1986). The 1944 Education Act: the 'art of the possible. In Harold L. Smith (ed.), *War and Social Change: British Society in the Second World War* (Manchester, Manchester University Press), 101–28.

Vincent, D. (1989). *Literacy and Popular Culture: England 1750–1914* (Cambridge, Cambridge University Press).

Walden, G. (1996). *We Should Know Better: solving the educational crisis* (London, Fourth Estate).

Wang Binhua (1995). Comparing HSCE in the People's Republic of China and GCSE in England. In Little, Wolf and Wang Gang (1995).

Wang Gang (1995). The Development of Public Educational Examinations in China from 1980. in Little, Wolf and Wang Gang (1995).

Wiliam, D. (1996a). Meanings and Consequences in Standard Setting. *Assessment in Education*, **3**(3), 287–307.

Wiliam, D. (1996b). Standards in examinations: a matter of trust? *The Curriculum Journal*, **7**(3), 293–306.

Wilmut, J. and Rose, J. (1989). *The Modular TVEI Scheme in Somerset: its concept, delivery and administration* (Report to the Training Agency of the Department of Employment, London).

Wolf, A. (1995). Competence Based Assessment (Buckingham, Open University Press).

Wolf, A. and Steedman, H. (1998). Basic Competence in Mathematics: Swedish and English 16 year olds. *Comparative Education*, **34**, 3.

Wood, R. (1991). *Assessment and Testing: A survey of research* (Cambridge, Cambridge University Press).

Young, M. (1958). *The Rise of the Meritocracy 1870–2033* (London, Penguin).