POLICY CHALLENGE

Date: September 1993

Centre for Policy Studies, 52 Rochester Row, London SW1P 1JU Tel: (071) 828 1176 Fax: (071) 828 7746

TESTING TIME

The Dearing Review and the future of the National Curriculum

John Marenbon

Price: £3.95

Note: nothing written here should be taken as representing the view of the Centre for Policy Studies, which never expresses a corporate opinion in its publications.

The Author

John Marenbon is a Fellow and Director of Studies in English at Trinity College, Cambridge. He was a member of the School Examinations and Assessment Council in 1992–3.

The Centre for Policy Studies never expresses a corporate view in any of its publications. Contributions are chosen for their independence of thought and cogency of argument.

ISBN 1 897969 12 0

© Centre for Policy Studies, September 1993

Printed in England by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk

Contents

1. Introduction 5

2. Sir Ron's Compromise 6

3. A Better Solution 14

1
Introduction

Sir Ron Dearing's *Interim Report** on his review of the National Curriculum has seemed to many observers a politic compromise between the demands of teachers and the requirements of the Government. But it fails to tackle the Curriculum's underlying problems, and its central proposals are confused and contradictory. Teacher assessment is to be given more importance in national reporting, but the criteria on which it rests will be simplified and loosened. Tests will be intended to moderate teacher assessment, although what each measures is not the same. League tables will be adjusted so that they no longer indicate how well children at different schools have in fact performed. Bureaucracy and prescription will be little reduced, and the Curriculum will continue to place severe and arbitrary restrictions on the freedoms of teachers, pupils and parents.

The Government must be bold. Piecemeal revision will not salvage a curriculum which is fundamentally rotten, because it does not allow for a diversity of views, teachers and pupils. It should be scrapped. A minimal curriculum should be substituted for it — a curriculum which occupies altogether less than half of school time, and is restricted to specifying those basic skills and areas of knowledge which every reasonable person would agree an education should provide.

---

* *The National Curriculum and its Assessment: an Interim Report* (York/London; National Curriculum Council/School Examinations and Assessment Council, 1993).

2
Sir Ron's Compromise

The chain of events leading to Sir Ron Dearing's review stretches back far further than the teachers' rebellion which made it necessary — back to the very beginnings of the National Curriculum itself. When, in 1987, Kenneth Baker announced that the Government would institute a National Curriculum, support was widespread. But many, especially among Conservatives, imagined that a National Curriculum would be a limited exercise, designed to ensure minimum standards in the most important school subjects: a minimal curriculum. While Mr Baker was in Mole Valley defending his seat at the General Election, his officials were not idle. On his return to the Department of Education, the Secretary of State was presented, not with a plan for a minimal curriculum, but rather a comprehensive scheme, occupying almost every moment of school time and prescribing for every aspect of ten disciplines: not only English, Mathematics and Science (the 'core' subjects), but also Technology, History, Geography, Music, Art and Physical Education (the 'foundation' subjects).

A strange triple alliance ensured that this ambitious scheme became Government policy and, finally, law. The first of the allies were the educationalists. They had long wished for a comprehensive National Curriculum, in order to enforce upon all teachers the fashionable methods and aims described by them as 'best practice'. The second of the allies were certain doctrinaire members of the Conservative Right, who had observed with distress the appearance in school syllabuses of subjects they found educationally and politically objectionable, such as Peace Studies. They wanted a curriculum, based on traditional school subjects, to take up as much of school time as possible, so that there would be no room left for anything else. The third ally was Mr Baker himself, anxious for a big Bill which could be smoothly implemented, because it pleased not only his Party, but also the teachers and educationalists.

This comprehensive Curriculum needed a general framework for assessment and testing. A group dominated by educationalists (the Task Group on Assessment and Testing) was entrusted with providing what became known as the TGAT model. TGAT set about its task in line with the accepted ideas of modern educational theory. Educationalists distinguish two types of assessment, 'diagnostic' and 'summative'. Diagnostic assessment is another name for the observation and reporting back of pupils' strengths and weaknesses which teachers have carried out from time immemorial — a process inseparable from the very business of teaching itself. Summative assessment, by contrast, provides a record of a

pupil's level of achievement at a given moment; and, from the last century until very recently, this assessment has usually been carried out through formal, written examinations. Modern educationalists tend to regard summative assessment with suspicion, unless it can be intimately linked to diagnostic assessment and carried out by the pupils' own teachers in the course of their ordinary teaching. They also treat diagnostic assessment as a far more regulated and formal process than it has traditionally been: less a matter of individual judgement by teachers than a measurement of progress against educationalists' criteria for what each subject should involve. The TGAT model was expertly designed to reflect and realize these presumptions.

Each subject was divided into a number of main aspects or 'Attainment Targets' — for instance, in English: Speaking and Listening; Reading; Writing; Spelling/Presentation; Handwriting/Presentation. In every subject except Art, Music and Physical Education, ten levels of attainment in every Attainment Target were then defined, each by means of one or more 'Statements of Attainment'. A Statement of Attainment is a brief description of some particular skill or type of understanding. For example, 5(c), one of five Statements of Attainment which define Level 5 in Reading (Attainment Target 2 for English) is that pupils should be able to 'show in discussion that they can recognise whether subject matter in non-literary and media texts is presented as fact or opinion.' Pupils who show that they can do this, and do what is asked by the other four Statements of Attainment, will have reached Level 5 in Reading. A pupil who cannot satisfy all five Statements of Attainment will not have reached that level.

Since (at least in theory) each school subject is described by its Statements of Attainment, grouped and arranged according to the order in which they should be fulfilled, they provide the ideal instrument for both its diagnostic and summative assessment. Suppose a child has satisfied four of the Statements of Attainment for Level 5 Reading, but not yet 5(c), then this one (discerning fact from opinion in non-literary texts) is precisely diagnosed as his point of weakness. Once he overcomes it and reaches level 5, then he can be judged, summatively, as having progressed further in reading than children at level 4, less far than those at level 6. Moreover, the TGAT model provides at least the illusion that it can be assessed by teachers in the course of their ordinary teaching in a way which does not vary from school to school. The more a subject is broken down into discrete units, the more plausible the idea that mastery of each can be gauged objectively by any person of competence and goodwill.

Unfortunately, this elaborate structure is deeply flawed both in theory and practice. Some of its problems are noted by Sir Ron (§4.5): for example, the unwieldiness of a system involving nearly 1,000 Statements of Attainment, and the atomization of teaching

and learning which it encourages. But it also has other and deeper defects. As a model of assessment, it is one dimensional. For a child to be good at a subject means, in its terms, simply that he can do at a younger age what most children will be able to do when they are older. Perhaps this model fits some subjects, such as mathematics; but for others, like English, it is entirely inappropriate. The talented eleven year-old may show a verve and imagination in his writing, a perceptiveness in literary appreciation, which neither pupils seven years older, nor even their teachers, can equal. Moreover, by describing in detail the steps by which the learning of each subject must proceed, the TGAT model removes all room for individual choice. It does not allow that different pupils learn in different ways and different orders; nor that different teachers may very properly have different views about what constitutes their subject. Most seriously of all, the TGAT model assumes that subjects can be properly described in a language external to them. Its architects appear to have had no notion that an understanding of what a subject is requires a deep knowledge of that subject; that it may be a matter for constant dispute among experts; and that it may admit of no neat formulation, perhaps of no formulation at all.

Despite these defects, the TGAT model was accepted with little opposition by the world of education. The burden of bureaucracy which the complex system would impose on teachers was yet to become evident. The model of assessment accorded so closely with what educationalists prized, and had taught teachers to prize, that its manifest deficiencies were ignored. Teachers might indeed have been expected to bridle at the prescriptiveness of the scheme. But the Curriculum Orders for each subject were drawn up in line with what was regarded as 'best practice' in the discipline. Those who had objections were isolated as extremists. Few so prize intellectual freedom that they will complain about prescriptions which they themselves are eager to follow.

Difficulties came from another quarter. Having set up the system, the Government began to have doubts. There was discontent among Conservatives over a Curriculum which seemed to give the force of law to progressive educational theories and attitudes. In particular, there was distress that elaborate schemes for teacher assessment had taken the place of the simple 'paper and pencil tests' which many of them had looked forward to welcoming. On this Kenneth Clarke, the new and resolute Secretary of State, determined to act. Assessment for 14-year-olds, first due to take place in some subjects in 1993, would, he decided, be mostly by formal, written tests. As it became clear to teachers of English — one of the subjects due to be tested — that the written tests (which most of them in any case disliked in principle) would not even accord with their dominant conception of what was important in their subject, disquiet turned into complaint, and complaint to outright rebellion. To make matters worse, the test papers had to be set so as to assess the

Statements of Attainment in the National Curriculum English Order, which had been drawn up with continuous assessment in mind; and so it became evident that the tests would be unsatisfactory in any terms and, in some cases, downright unfair. Meanwhile, teachers in every subject became ever more painfully aware of the administrative workload which the National Curriculum was imposing on them. Here was a simple and legitimate grievance, an issue which — unlike their more technical dispute about methods and aims of assessment — teachers could use to gain public support. The lines of battle were drawn, and perhaps there was little which Mr Clarke's unfortunate successor could do, by way of compromise or standing firm, to prevent the painful conflict which ensued.

It was at this point — shortly before thousands of pupils were due not to take their tests — that Sir Ron began his review. He faced, therefore, an immediate political imperative: to discover some compromise in the two main areas of dispute between teachers and the Government — administrative workload and methods of assessment (formal tests or continuous teacher assessment). But Sir Ron also needed to tackle the underlying problems of the National Curriculum. Unless they are resolved, any changes he makes can do little lasting good and may well do harm, by perpetuating a scheme which is fundamentally misconceived. These problems are (to summarize):

(1) The all-embracing, prescriptive character of the Curriculum, which prevents teachers from teaching according to their differing views and interests, and according to the particular needs of different children.

(2) The grave educational defects of the TGAT model.

(3) The incompatibility between the TGAT model and formal, written tests.

Sir Ron's *Interim Report* has now been published. How well does it deal with these difficulties?

Sir Ron's two central proposals are compromises. The main structure of the National Curriculum will remain and, even if the 10-point scale is abandoned (a course which, on balance, he does not seem to recommend), it will be replaced by a different system of levelled criteria.* But the number of Statements of Attainment will be greatly reduced and, in some cases, individual Statements of Attainment will be made looser and wider (§§1.7 (ii); 4.9). Formal tests, too, will remain, although they will be shorter. But the results of teacher assessment will be reported separately and given equal standing to those of the tests (§§1.18; 5.17, 28).

---

* These would be attached to the ends of the different 'Key Stages' (7, 11, 14 and 16). Sir Ron envisages trying to establish, for each subject at each stage, criterion-referenced statements of the kinds of knowledge, understanding and skills which average pupils, and those half-way up above, and half-way down below, the average should show (§§4.24–9).

Whether these compromises answer Sir Ron's political imperative successfully, the events of the coming months will tell. Fewer Statements of Attainment will mean less bureaucracy, but not all that much less; and the renewed emphasis on continuous assessment might, in the long run, make for a heavier burden of administration than if the National Curriculum had remained unaltered. Teachers in some subjects will accept an equal weighting between tests and their own continuous assessment; in others, where course work has become almost a cult, they will not surrender so easily. Sir Ron himself seems rather unsure about the differing functions of formal tests and teacher assessment. In the summary to his report (§§1.12–15), he suggests that teacher assessment should be mainly diagnostic in purpose, whilst formal tests should have 'a key role' in providing summative assessment. This view, however, is plainly at odds with the decision that teacher assessment should carry an equal weight to test results in national reporting. How will Sir Ron argue for his proposal, when he himself has no clear idea of what it is supposed to achieve?

More seriously, Sir Ron's compromise, whether effective politically or not, fails to resolve the serious, underlying problems of the National Curriculum. The first of these, listed above, is the Curriculum's prescriptiveness. It has been widely reported that Sir Ron's plans will make the National Curriculum far less prescriptive and give teachers much greater scope to exercise their professional judgement. Yet the changes in this direction proposed in the *Interim Report* for pupils aged from 5 to 14 are minor. A little time will be liberated by making some of the material in the programmes of study for foundation subjects non-statutory (although this non-statutory material will continue to be drawn up by the Government's official advisers). Since the National Curriculum and religious education will still occupy between nine tenths and three quarters of school time — hardly less than it does at present — this concession is slight. The same eleven subjects will continue to be obligatory for all children up to 14 (§§3.16–29). Breadth and balance — worthy objectives indeed for a curriculum — will continue to be interpreted to mean uniformity, and different pupils will not be given the chance to develop their differing talents to the full.

Moreover, prescription in the core subjects (English, Mathematics, Science) will be reduced, if at all, even less than elsewhere. In his summary, Sir Ron remarks that 'the central importance of the core subjects ... means that the statutory core in these subjects will be larger than that of the other subjects' (§1.7(i)). This is a *non sequitur*. The importance of a subject should certainly determine the amount of time given to it. But the extent to which the content of a subject is fixed through central prescription should depend, rather, on the degree to which there is general agreement on what study of the subject involves. In the case of at least one of the core subjects, English, the lack of such agreement is all too

evident. Almost everyone accepts that, in English lessons, children should be taught to read accurately and write correctly; and also that this teaching of basic literacy, although essential, is only a limited part of the subject. But there are a thousand different views as to what else should constitute English, and no arguments for one or another of them which are so compelling that any reasonable person would accept them.

Towards pupils aged 14–16, Sir Ron's attitude is different. Here he is willing to consider allowing a good deal more choice of subject than the National Curriculum permits at the moment. In particular, he suggests that some pupils might benefit if they could choose a more technically or vocationally directed course at 14 (§§3.2–8, 27–28). This suggestion is, in itself, a wise one, but Sir Ron attaches it to another which has a very sinister implication. Rather than thinking about education in terms of ages 14–16 (Key Stage 4 of the present National Curriculum), Sir Ron proposes that we might, as in some other European countries, treat the years from 14–19 as a single unit. But to do this would open the way for the rules and restrictions of the National Curriculum to creep beyond the years of compulsory schooling. Under the cover of relaxing the Curriculum for 14–16 year-olds, is Sir Ron trying to extend prescription to 16–19 year-olds as well?

The other two problems underlying the National Curriculum concern the TGAT model: its educational disadvantages and its unsuitability as a basis for formal tests. These difficulties, as has been explained above, are not ones which arise from the particular structure of the TGAT model, but rather are intrinsic to any all-embracing criterion-referenced system. Sir Ron proposes changes — possibly even far-reaching changes — to the TGAT model, but he intends to keep a system which is comprehensive and involves many explicitly stated criteria for achievement. His modifications will not, therefore, remove the most serious problems in the present scheme: the one-dimensional approach to assessment; the attempt to fix, delimit and describe what cannot be fixed, delimited or described; the irresolvable tension for those devising test papers between the need to measure fulfilment of Statements of Attainment and the demands of good question-setting, between the necessary narrowness of test questions (especially if they are to be marked according to a rigid scheme) and the breadth of criteria designed to describe the whole of a subject. Indeed, there is every possibility that, if whole subjects are described in a smaller number of Statements of Attainment, these problems will only be exacerbated.

Not only does Sir Ron's compromise fail to resolve the fundamental problems of the National Curriculum, it will also create a new difficulty all of its own. As the architects of the National Curriculum understood, if teacher assessment is to give even the illusion of providing a standard measure of achievement from school to school and county to county, it must be conducted by using a set of very precise criteria. Perhaps in some subjects these

criteria are not really very precise at all; but it is certain that, if the criteria are broadened and made looser, they will become even less precise. Sir Ron talks of developing assessment criteria in certain areas 'which recognise that performance depends upon a complex of knowledge, understanding and skill which is not easy to divide into discrete elements' and which would require teachers 'to exercise a greater degree of judgement' than at present (§4.9). Fine words! But how can criteria which are so loose and dependent on interpretation function as criteria at all? Sir Ron's changes will make the National Curriculum less effective than it is now as an instrument for ensuring some degree of standardization in teacher assessment. Yet, at the very same time, Sir Ron intends to raise the status of teacher assessment in national reporting. Are not these moves — the two main elements of his compromise — at odds with one another? It is as if a man were at once greatly to weaken the structure of his edifice and yet insist that it bear a far greater weight than it carries, insecurely, at the present. This cannot be sensible policy.

Sir Ron himself seems to detect a problem. He recognizes that 'effective moderation' of teacher assessment is necessary, especially given his loosening and simplifying of the Statements of Attainment. But he imagines that this moderation will be supplied by the tests (§§5.14, 17(iii)). He does not consider the obvious objection that, since what continuous teacher assessment measures will often be different from what is gauged by the written tests, the one cannot be moderated by the other. It is always at least arguable that a pupil's performance on a given day misrepresents his work over the year; and continuous assessment will embrace aspects of many subjects which tests do not. In English, for instance, continuous assessment will take into account speaking and listening, and also breadth of reading, It will not be surprising, therefore, if children gain different levels from teacher assessment than they do in the tests.

Why not, then, abandon formal tests entirely — a move which would certainly be welcomed by most teachers? The answer is simple. To do so would be disastrous; but not for the reason which many imagine. The problem with teacher assessment is not that it cannot, at least in theory, provide trustworthy, standard gradings, but that it can do so only at the cost of ruining the education which it measures, by failing to respect the limitations intrinsic to any type of summative assessment.

No system for grading pupils' achievements, whether by teacher assessment or formal tests, can be of more than limited value. The reasons are various. First, many subjects involve abilities and modes of understanding which are, in the strict sense, not comparable with each other: that is to say, if a certain child has gained one set of them to a high degree, and a different child another set to an equally high degree, there is no way of establishing which child is better at the subject. Second, many of the gains which children make from

good teaching are not immediate ones. An excellent teacher of 7 year-olds may see the full results of his work only when his former pupils flourish at university; and the deepest benefits of education are evident only through the whole course of a lifetime. Third, it is the job of schools not merely to impart knowledge, but also to instil a love of it. A devotion to learning for its own sake is not, however, something which can be measured or graded.

None of these considerations is intended to undermine the importance — limited but definite — of summative assessment. Clear gradings are useful as indicators to pupils themselves, their parents and, for older children, to employers and selectors too (although it would be a foolish company or university which chose its entrants solely on the basis of gradings made by teachers or examiners). Summative assessment should be regarded as a necessary evil, detached from the central aims of education and, quite often, at variance with them.

When summative assessment takes place through formal, end-of-course examinations, it can be kept in its place. The tests can be largely ignored until the last moment, whilst teachers and pupils pursue their central aims of teaching and learning. By contrast, continuous teacher assessment is, as its name indicates, continuous. Pupils are on test all the time. Teachers must constantly turn their minds from teaching to assessing — not the informal type of assessment which is part of good teaching, but assessment designed to assign a definite grade or level to each child. Not only is the proper work of the classroom disturbed. Inevitably, the relations between teachers and pupils are corrupted. Pupils cannot afford to stand up to a teacher or fall out with him, since they cannot look to the results of an independent test for some vindication of their abilities. Teachers, whose schools will be judged by the grades they give, cannot afford to mark their pupils harshly, even if this is what they believe they deserve. Moreover, as has been argued above, teacher assessment can claim to provide standard gradings only if it is based on an elaborate scheme of fixed criteria. But such a scheme restricts freedom and distorts the nature of many of the subjects it purports to describe.

In brief, neither tests nor teacher assessment can avoid the defects intrinsic to all summative assessment. By confusing teaching and testing, a system of teacher assessment allows these defects to invade every part of education. A system of formal tests, by contrast, is limited and capable of recognizing its own limitations. Although there will always be some bad teachers who teach to the test, the system does not make the mentality of grading and measuring omnipresent and inescapable.

3
A Better Solution

The National Curriculum in its present form is too deeply flawed for piecemeal revision. Besides the very serious problems which Sir Ron has failed to tackle, there is an even more fundamental objection to the whole conception on which the present Curriculum is based. There are very many different views and arguments about what is required to make up the whole of a good education. Despite the single-mindedness with which some of these are advocated, none is so compelling that it demands the assent of all who consider it open-mindedly. A comprehensive curriculum, which imposes a single view of education on everybody, arbitrarily limits freedom in an area where teachers, parents and pupils have every justification for exercising it. No government should impose such a curriculum on a free society — least of all a Conservative Government which has made choice and diversity the slogans of its educational policies.

The Government should begin afresh and institute a minimal curriculum. There are certain basic skills and areas of knowledge which almost everyone, whatever his views about the aims of education as a whole, would agree are necessary as parts — but only parts — of a good education. Hardly anyone would wish a child to leave school unless he could read well, write Standard English correctly, was proficient in simple arithmetic and geometry and knew the most elementary scientific facts and principles. Most people would agree, too, that children should learn at least one, if not two, foreign languages. Here is the basis for a minimal curriculum. Its specification should be carefully designed so that it does not stray beyond these fairly narrow areas of broad agreement. It should not, by any means, discuss or prescribe methods of teaching. A minimal curriculum of this sort would take up more time in the early years, when most basic skills have to be mastered; but even then it would leave about half the timetable free. It would intrude less and less upon the work of all but the weakest pupils as they entered secondary school; and the average child would have finished it entirely before he was 14. This curriculum would serve limited but valuable purposes. It would identify children whose education is failing because they are not mastering even basic skills and knowledge, and expose schools where this type of failure is common. It would ensure that, in their anxiety to fulfil their conception of the wider aims of education, teachers do not neglect the foundations of learning. And, by limiting prescription to a commonly agreed core, it would show teachers that, in every area beyond this, they are free to teach what they, their pupils and their parents wish, no matter what pressures are put

on them by education authorities, inspectors or other officials to conform to what is regarded as 'best practice'.

There would be nothing in this minimal curriculum which could not be tested by simple, formal tests, marked according to rigid mark-schemes. In order to allow for the great differences in ability between children, whilst avoiding a complicated apparatus of tiering, these tests should consist of graded papers to be taken when children are ready for them. An average child might be expected to pass the lowest test at 7 or 8, the highest one at 13 or 14. The results of these tests would not reveal — and would not be taken to reveal — how good or bad children are in each subject as a whole, but merely whether they had acquired certain basic skills and areas of knowledge. None the less, since the teaching of fundamentals is so important a part of the work of primary schools, one measure of these schools' success could reasonably be provided by the test levels reached by their pupils.

This point raises another contentious question: should the Government publish league tables of schools' results? Most teachers believe not; but the Government is determined to persist with those based on pupils' performance at 11 and 16. Sir Ron has tried to negotiate a compromise in this area, too, by considering whether league tables can be devised which indicate 'value-added' (Annex 5). 'Value-added' is, in Sir Ron's usage, an amorphous term. He begins by talking of value-added tables as if their purpose were simply to indicate, not the absolute levels reached, but the number of levels by which pupils had risen during a period of their schooling. In the course of the discussion, however, it becomes clear that they would also make adjustments based on the gender balance of schools and on children's home backgrounds, and that they would try to take into account factors 'less tangible' than academic achievement. Both these ideas have very worrying implications. If, as a matter of statistical fact, children of one sex or from certain backgrounds do better or worse in a given subject, should this be accepted and institutionalized, as if their sex or their background inevitably made the children less able? And, laudable though it is to recognize all aspects of a school's achievement, when academic performance is proving so hard to assess accurately, it is hard to accept with equanimity the suggestion that 'less tangible' aspects of a school should also be made subject to official measurement.

A different compromise needs to be made. Teachers should realize that, by objecting to league tables, they are asking unreasonably for information, which the public would find valuable, to be withheld from it. For its part, the Government should acknowledge that the publication of league tables is not its business. It should restrict itself to ensuring that full information about academic performance is made freely available by every school. Those who, following public demand, produce tables should decide how they should present them. Should they wish to use this or that measure of 'value-added', by all means allow them to

do so. But it is hard to imagine that any tables would be read with such eagerness as those offering the plain data.

Here, as everywhere in the organization of the National Curriculum and of education in general, the Government would serve its best interests, and those of electors, by carefully regarding and respecting its own limitations. Ambitious schemes to raise the educational standards of the nation make for good political rhetoric, but not for good teaching.