Review: National Assessment and Testing: The TGAT Report
Author(s): Keith Kimberley, Ian Hextall, Harry Torrance and Bob Moon
Review by: Keith Kimberley, Ian Hextall, Harry Torrance and Bob Moon
Source: *British Journal of Sociology of Education*, Vol. 10, No. 2 (1989), pp. 233-251
Published by: Taylor & Francis, Ltd.
Stable URL: http://www.jstor.org/stable/1392642
Accessed: 15-04-2015 10:17 UTC

REVIEW SYMPOSIUM

National Assessment and Testing: the TGAT Report

Reviewed by Keith Kimberley, Ian Hextall, Harry Torrance & Bob Moon

National Curriculum: Task Group on Assessment and Testing (1987)
A Report
London, DES

National Curriculum: Task Group on Assessment and Testing (1988)
Three Supplementary Reports
London, DES

Looked at in terms of the overall intentions of the Government, the 1988 Education Act provides a series of inter-related devices for exerting control over the State school system. Within this outer framework, the National Curriculum will provide a detailed specification against which the performance of students can be measured, with the results of individual students being aggregated to provide information for the wider functions of the system, enabling market forces to be brought into play as parents pick and choose between schools.

Since the assessment of students is central to these arrangements, the Report of the Task Group on Assessment and Testing (DES, 1987) and the subsequent Supplementary Reports (DES, 1988) have significance, not just with respect to checking on the delivery of the National Curriculum to students and measuring their performance against its criteria, but also for the functioning of the whole enterprise. Whatever it may make possible in terms of student learning and entitlement, it will provide also the key accounting unit, on the basis of which the competence of teachers, schools, and LEAs will be assessed.

The Task Group on Assessment and Testing (TGAT) have thus had to operate within conflicting parameters, which they were not magically able to resolve. On the one hand, they were required to design a system to give information to students (so they can know their strengths and weaknesses); teachers (so they can provide appropriate learning experiences); and parents (so they can relate their child's progress to national performance) (DES, 1987, para. 126). On the other hand, they were expected to suggest ways of making available statistical data, in reports produced by schools and authenticated by LEAs (DES, 1987, para. 134), for parents, prospective parents, governors, education officers, and a wider public. With respect to the former, they have adopted a progressive rhetoric which emphasises positive achievement, flexibility, and care for the learner's

needs. With respect to the latter, they use the language of reporting, aggregating, and acountability derived from the overarching system.

One way in which the intractability of TGAT's task shows through is in their insistence that the package of proposals which they have devised must be adopted as a whole. The First Supplementary Report (DES, 1988a) concludes with an attempt to head off those critics of their main report who have claimed that less complex arrangements than those proposed are needed for the purposes of assessment at 7, 11 and 14. TGAT argues that it is, of course, possible to conceive of simpler systems but that no alternative has been put forward which demonstrably achieves all of the criteria in their remit: criterion-referencing to identify specific achievement; progression in learning; a formative system to guide the next steps; confidence in the results; and consequent improvement in teaching and learning (DES, 1988a). They challenge others including, we may assume, those who discuss these matters with the Prime Minister, to suggest which of these criteria should be abandoned: "the system we have designed stands or falls as a whole", adding that their "prime concern has been to enhance the professional work of teachers so as to improve the education of our children" (DES, 1988a, para. 26).

TGAT was in an impossible situation. If the assessment arrangements were to fulfil the requirements of the Secretary of State, first they had to be 'simple', 'clear' and 'cost-effective' (DES, 1988d, p. 69) and secondly must balance the promotion of learning with the need for summative information for "publicising and evaluating the work of the education service" (DES, 1987, appendix B, para. 6). Within the grip of these constraints, it is doubtful whether they could do more than attempt to humanise the assessment and testing operation. This they have attempted to do by building from wholistic principles; by insisting that all assessment should take place in normal learning contexts; by asserting that their system will encourage all children towards positive achievements; and by resisting pressure to use only those modes of assessment which are relatively inexpensive. However, despite this substantial and positive programme, the system within which they are working remains firmly tied to functions which are focused elsewhere than on the interests of the student.

Attainment Levels

At the heart of the TGAT scheme lie the *attainment levels* which are the TGAT response to the Secretary of State's demand for differentiation at each Key Stage:

> the measurement and recording at a range of different levels of positive achievement in reasonably discrete elements within each common target area. (DES, 1987, appendix B, para. 12)

The TGAT interpretation of this instruction has been to propose that

> the ways in which the criteria and scales are set up and used should relate to the expected routes of educational development, giving some continuity to a pupil's assessment at different ages: the assessments should relate to *progression*. (DES, 1987, para. 5)

and to elucidate this concept further they tell us that

> it is *not* necessary to assume that the progression defined represents

> some inescapable order in the way children learn, or some sequence of difficulty. Both of those factors *may* apply, but the sequence of learning may also be the result of choices, *for whatever reason*, which those formulating and operating the curriculum may recommend in the light of teaching experience. (ibid., my emphases)

I think that I may not be alone in finding the resonances of this statement disturbing. For example, the English Working Group in *English for Ages 5–11* (DES, 1988d) has chosen a rather different emphasis, arguing that "if proposals for assessment are to have coherence, they must be based on a theory of difficulty" (para. 10.15) and, in their responses to *English for ages 5–11*, the London Association for the Teaching of English draws attention to the dangers of establishing arbitrary hierarchies which will set "a supposed order of development to which teachers will teach and against which children will be tested", with a possible consequent lowering of student and teacher expectations (LATE, 1989, p. 2).

Such fears are confirmed by the mechanism for establishing the attainment levels which is described in the First of the Supplementary Reports:

> Initially the norms now expected for different ages will be used in helping to identify criteria appropriate for the system of ten levels; but, once devised, the system will rest on the levels and criteria alone. (DES, 1988a, para. 4)

and then in the Second Supplementary Report:

> it does not depend on empirical evidence of a particular linear, or other, pattern of learning for its initial construction, although the definitions of the levels may need to be reviewed in individual cases in the light of information about the actual distribution of pupils' performance when the national curriculum and assessment system are in operation. (DES, 1988b, para. 7)

If I understand this correctly, TGAT are willing to see levels established without much recourse to empirical evidence or developmental theories, with some modifications being made later (probably on the basis of the standard assessment tasks (SATS)). Through the subject working groups, versions of current norms are to be arranged in hierarchies which will, in due course, be revised and legitimated using information gathered by putting the scheme into operation. TGAT make a virtue of an initial "rough and ready" approach (DES, 1988b, para. 12) to the establishment of the attainment levels, and play down their subsequent role in the total system.

Effects on Students

Looked at from the point of view of the students, and especially those who have been at a disadvantage or discriminated against within the present system, the new arrangements have some positive aspects. The curriculum will be made visible and it will be possible to analyse the view of the world that it represents. Theoretically, this should mean that a nationalistic, racist, or sexist curriculum will not be written, or, if proposed, can be challenged. It should be possible

through public pressure to ensure that the necessary revision and rewriting takes place. A national curriculum is, by nature, visible.

There should also be guaranteed routes of access for all students to the whole curriculum. Theoretically, this could mean an end to complaints from ethnic minority parents concerning inequalities in access to high status knowledge. These views are held strongly, particularly by those who had, on arrival as immigrants, high hopes for their children in British schools only to find their children often being offered watered-down, low-prestige courses and discouraged from pursuing academic subjects. A national curriculum entitlement, together with reports on individual students' progress at important points in their school careers, could in this context help to redress some long-term grievances.

The desire for a quality education for ethnic minority students may not, however, be met in practice. The establishment of a national curriculum also can be seen as an attempt to place the Government's views of British society centrally in the curriculum and to roll back any space for curriculum content and activities thought by the Government to be subversive or potentially critical. One function of this move is to shift attention away from locally inspired curriculum initiatives, and particularly from those which attempt to tackle issues of gender, race and class.

Further, while the new system may make it possible to see at a glance where a particular student is in relation to 'national performance' (DES, 1987, para. 126), it is by no means clear that this will operate in the best interests of all children—or even if that is what the Government intends. Many students may, indeed, be encouraged by having targets to aim for but on the underside of the system there will be students who will be discouraged by achieving only low levels of attainment. All those students whose school careers are marked by uneven development, with periods of slow progress and sometimes regression inerspersed by periods of sudden, dramatic activity (often sparked by contact with particular teachers or particular content), may be harmed by a system in which you find your attainments so precisely defined at key points in your school career. A greater discouragement still would be to find yourself labelled by teachers, or grouped, in terms of the levels you have achieved at 7, 11 or 14.

Students whose first language is not English could experience some long-term benefits from the new system if the results from individual schools demonstrate the efficacy of good, bilingual provision in enabling students to reach high levels of achievement. TGAT has also interestingly proposed that, at 7, students should be assessed in other subject areas than English through their first languages "wherever practicable and necessary" (DES, 1987, para. 53), which, if implemented, would be a valuable step towards a nationwide recognition of the importance in children's learning of their first languages.

For those who will be assessed in English—the vast majority since only those who have so little English as to "render the assessment unworkable" (ibid.) are to be exempted—TGAT attempts to provide reassurance that low levels of performance because of language difficulties "would be no reflection on a pupil's general ability but merely an indication that a pupil needed special help with language skills" (ibid.). It remains to be seen how the students themselves will feel and whether they will believe that their results are "no reflection on their general ability" (ibid.).

The English Working Group has followed TGAT's lead in presenting a positive

view of bilingualism and, by endorsing the ILEA Primary Language Record (DES, 1988d, para. 9.20), has drawn attention to ways of acknowledging and charting the progress made by bilingual students. Such approaches rely on integrating an understanding of what it is to be a bilingual student with general descriptions of attainment and are in keeping with TGAT's comments on the communication of information to parents, who play an important role in the ILEA scheme.

TGAT suggest that, in assessing a student

> Most straightforwardly one could simply describe what a pupil understands, knows, and can do in terms similar to those that will be used by the subject groups to define attainment targets. (DES, 1988, para. 196)

I take this sentence to be a very important assertion of the value of description in assessment and to mean that teachers could employ generally acceptable, common-sense definitions in non-technical language for many purposes in assessing children's attainments. Unfortunately, TGAT also are committed to converting and reducing the descriptions into a marking scale for reasons for comparability and accountability (DES, 1987, para. 96). In a system which simultaneously faces two ways we may speculate as to whether detailed descriptions or grades and scales will prove the stronger and whose interests will be served by the outcome.

Another set of possible consequences for students and their parents stem from the effects that the new arrangements may have on schools. Schools may become very preoccupied with distributions of aggregated, numerical scores. Determination to boost their public image may have significant effects on school organisation and ethos. For example, there may be disincentives for secondary schools to take on students with low levels of attainment, who they think may depress their ratings or be costly to provide for.

There is no guarantee that, because of the new system, teachers who currently have low expectations of ethnic minority students will begin to see them in a new light. It is, of course, likely that there may be benefits for students if poor teachers are identified and replaced, harsh though this sounds. Teachers certainly will have incentives to try to get students from one level to another. However, some may be tempted to employ deficit explanations of under-achievement, blaming the student for his, or her, failure to shape up to the requirements of the national curriculum. One organisational form which this could take would be an increase in streaming and setting, together with their potential for institutional racism. A less visible, but equally damaging consequence would be a return to the low expectations of many students and the narrow mental set which streaming and setting often involve.

Indeed, there is insufficient discussion in the TGAT Reports of the possible effect of the proposals on pedagogy. It is optimistically assumed that even the SATS will become part of a day-to-day practice and will not stand out from normal learning contexts (DES, 1987, para. 16) and that good teachers will be able to continue to teach flexibly and imaginatively. There are other less attractive scenarios. Teachers may begin to adopt the kind of standardised teaching formats which they feel are necessary if students are going to be able to handle the SATS. There may also be a massive increase in standard textbooks, perhaps on the North American pattern. Neither of these possibilities are likely to raise levels of expectation. Further, where the new assessment arrangements require a shift in pedagogy from one which is predominantly student-centred to one which is

dominated by defined subject content there may be far-reaching consequences in terms of student alienation.

In conclusion, the role of assessment within the framework of the 1988 Education Act is to provide standardised, nationally comparable results which can be made available for inspection by all interested parties; students, parents, teachers, headteachers, LEAs, local communities, and central government (DES, 1987, paras 45, 96, 125–138). Publication of results, together with open enrolment, local financial management, and the potential for opting out, make up the rest of the package.

For the whole process to have credibility, it will have to draw on such consensus as there is in England and Wales with respect to a common curriculum and, at first glance, there would appear to be some general ageement about the possible beneficial effects of national curriculum guidelines. As the process of curriculum specification and test development gets under way, some sense of the potential for consensus has been shown by the fact that the subject associations for the core subjects seem to have been able to accept many of the proposals from the Working Groups (though we should note that they have strongly opposed alterations subsequently made by the Secretary of State). Indeed, many school students and their parents may take pleasure from the discovery that the new curriculum framework offers wider opportunities than they were offered before.

This said, a second glance reveals a failure to take the opportunities offered in the making of a national curriculum. The Swann Report called on the Secretary of State to bring about a fundamental change in the education of every school student which would

> give every youngster the knowledge, understanding and skills to function effectively as an individual, as a citizen of the wider national society in which he [*sic*] lives and the interdependent world community of which he is also a member. (DES, 1985, p. 319)

This message has been acknowledged, to some degree, by the core subject working groups (for example, *Science for Ages 11–16*, paras 4.42, 7.12–7.17 and *English for Ages 5–11*, para. 6.3) but there is evidence of little more than a token nod towards the exhortation from Swann by the present incumbent of the office of Secretary of State for Education, whose public utterances have asserted the importance of English (*pace* Wales) traditions, English history, English literature, English culture, and standard English (Baker, 1986, 1989).

TGAT has not, however, like the working groups, had to engage with what view of society and the world should be built into a national curriculum. Instead, it has had the luxury of being limited by its terms of reference to "practical considerations" (DES, 1987, appendix A) and, with respect to their task, TGAT seem confident that they have devised a system of assessment which can meet all the challenges made to it. This confidence manifests itself in a persistent optimism about the way in which the whole set of assessment measures will be taken on by students, teachers, schools, LEAs, and a wider public and the safeguards they have suggested to prevent any narrowing of their proposals.

There is plenty in the TGAT proposals which, admirably, seeks to take the Secretary of State at his word and create a climate in which assessment is seen as integral to learning. But even taken as a whole, as they intend, it seems that some students may find the constraints of the system less than helpful. In time, the

utopian TGAT view of teachers working on the 10 levels "in collaboration with their pupils" (DES, 1988a, para. 20) may give way to harsher realities.

Correspondence: Keith Kimberley, University of London Institute of Education, 20 Bedford Way, London WC1 0AP, United Kingdom.

REFERENCES

BAKER, K. (1986) *Speech to students at Pangbourne College, Berkshire*, DES Press Release.

BAKER, K. (1989) *Terms of Reference for National Curriculum History Working Group*, DES Press Release.

DEPARTMENT OF EDUCATION AND SCIENCE (1987) *National Curriculum: Task Group on Assessment and Testing: a report* (London, DES).

DEPARTMENT OF EDUCATION AND SCIENCE (1988a, b, c) *National Curriculum: Task Group on Assessment and Testing: three suplementary reports* (London, DES).

DEPARTMENT OF EDUCATION AND SCIENCE (1988d) *English for Ages 5–11:* Proposals of the Secretary of State for Education and Science and the Secretary of State for Wales (London, DES).

LONDON ASSOCIATION FOR THE TEACHING OF ENGLISH (1989) *Responses to English for Ages 5–11* (London, LATE).

In a review as brief as this it is only possible to *assert* contextual points which clearly need to be argued more thoroughly. First, that no specific policy development, whether in the field of education, health, housing or defence can be adequately grasped outside the material, political and social context in which it is located [1]. Secondly, the analytic fragmentation of policy into specific elements may lose the essential interrelatedness which provides its rationale and dynamic. Thus it is apparent that the proposals on national assessment/testing cannot be disentangled from other aspects of the Thatcherite programme of educational restructuring.

Assesssment, curriculum, open-enrolment, grant-maintained status, city technology colleges (CTCs), local financial management, the delegitimation of LEAs, not to mention the destruction of the ILEA, are ideologically and programmatically interconnected. By extension, so are attacks on teacher unions, the restructuring of post-school education, and the 'stage whispers' regarding teacher education [2]. In conjunction with policy initiatives in other sectors they constitute an avowed, overt project. In essence this entails a destruction of what has come to be called the post-1944 'liberal-democratic settlement', and the exclusion of socialism from the agendas of current and future political debate. Within this overall pattern of development there are, however, some quite specific points which can be made about the national assessment proposals.

The overall tenor of the TGAT report is such as to provide a legitimated basis upon which it is possible to *differentiate* between LEAs, schools, classes of pupils, pupils themselves and, by extrapolation, teachers. The very sophistication and subtlety of the report forms a sweet-smelling fog which hides the swamp of inequality to which it leads. In the context of open-enrolment, opting-out, and CTCs, it is but a short step from differentiation (which in other spheres is called 'targetting') to selection. The prevailing assumption of hierarchy can be clearly

seen in Fig. 1. In this diagram the discreited 'bell-shaped curve' is turned on its side and ironed-out into a much more congenial rising, straight-line graph.

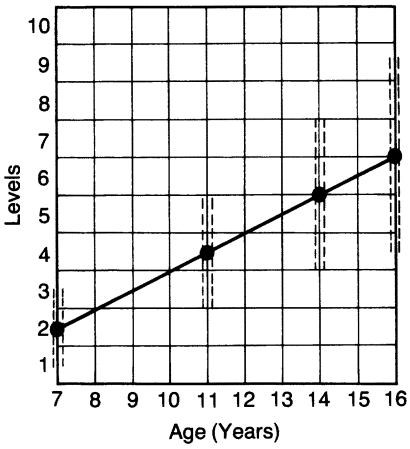

FIG. 1. Sequence of pupil achievement of levels between ages 7 and 16 (TGAT, para. 104).

As schooling is being restructured to create a 'disparity of esteem' and inequality of resources, competitive self-interest becomes the 'rational' response of parents, governors, pupils and teachers. Thus the whole TGAT report assembles a complex repertoire of assessment proposals, the outcome of which is to discriminate, in both senses of the word. In so doing it promises 'results' which can form the concrete basis on which 'informed choices' can be made. 'Privatisation' now takes on an internalised, ideological form to complement its material expression in the market sphere.

Structurally this links directly to social divisions. In the face of such striving individualism even the mildest forms of social redistribution or compensation are derided as expressions of 'dependency culture'. It is not eough to:

> . . . recommend that assessment tasks be reviewed regularly for evidence of bias, particularly in respect of race and gender. (TGAT, para. 52)

It is not just that 'class' does not exist, nor that 'sexuality' is invisible. Research has indicated that the very context, structure and process of such modes of assessment are inherently (perchance, intentionally) oppressive and discriminatory [3].

These characteristics are not vagaries of oversight, nor limitations of technique but specific, structural features of precisely the kinds of assessment modes to which TGAT is pointing. The measured balanced tone of the report disguises the process of objectification and reification which is underway. In an earlier article I cited the following quotation from the Runnymede Trust and Radical Statistics Group which bears repetition:

> Official statistics tend to be regarded as particularly authoritative and objective sources of data. This view is supported by the Government Statistical Service, which portrays itself as a neutral fact-finding agency, as a kind of statistical 'camera' used by the government to provide information to help run social affairs more effectively. . . . This view of

> official statistics is, however, not the real situation at all. If taken a little further the snapshot analogy may itself be used to challenge this view. We may go on to ask whether the nature of the numerical picture is dependent in any way upon who takes the picture, the particular instruments they use or the requirements of those who commission the picture in the first place. To pose these questions is to raise the whole issue of whose picture of society is reflected in these data. [4]

In a journal devoted to Sociology of Education it is also worth recalling what Pierre Boudieu argued many years ago:

> In fact, to penalize the underprivileged and favour the most privileged, the school has only to neglect, in its teaching methods and techniques and its criteria when making academic judgements, to take into account the cultural inequalities between children of different social classes. In other words, by treating all pupils, however unequal they may be in reality, as equal in rights and duties, the educational system is led to give its *de facto* sanction to initial cultural inequalities. [5]

In the introduction to the TGAT report it is clearly stated that it is important to adopt:

> . . . comprehensible language for communicating the extent of those achievements to pupils, their parents and teachers, and to the wider community, so that everyone involved can take informed decisons [choice, selection, resource allocation] about future action. (TGAT, para. 2, my additions in brackets)

This would be a quite proper and even laudable intention in other contextual circumstances. However, the Report itself belies a serious commitment to this intention. Both it and the summary version are characterised by a technicism of form and content which is mystifying and alienating and yet carries a high 'plausibility loading'. Paragraph 6 of the first Supplementary Report, intended to clarify some of the key arguments, amply illustrates the point:

> 6. Whilst profile components are clusters of attainment targets, it is possible that where some profile components are analytic and others synthetic in the same subject, the same attainment targets may feature in both (e.g. first in knowledge, and then in use of that knowledge). We also envisage that a target may be defined as a single criterion (e.g. can use a clinical thermometer) or as a linked strand—or sequence—of criteria (e.g. measurement of temperature). The former applies at only one level, while the latter can be applied through detailed criteria at several levels. Either usage is legitimate but the two should not be confused. We think that the second usage—regarding an attainment target as a strand of related criteria defined at several levels—will be the more useful in securing economy and clarity for communication purposes.

Once again the Radical Statistics Group illuminates the issue with force and clarity:

> A full assessment of an educational research study must take account of the fact that its ultimate use is generally to persuade people, and to produce support for policies and practices. [6]

In educational terms what used to be 'gentling the masses' has now shifted into 'the engineering of consent' with all the appropriated Gramscian overtones so beloved of the New Right. As the NCCL have recently argued, civil liberties can be eroded not only by their explicit removal but also by so corrupting concepts of 'freedom', 'liberty', 'equality' and 'rights' as to render them obscure, confusing and distorted [7]. A direct expression of this can be seen in the fact that none of the Educational Reform Act documentation available to parents has been distributed in any community languages other than English or Welsh.

The ideological coercion which underpins such 'consent' is evident in the highly centralised manner in which the assessment and curriculum policy is to be formulated and delivered. Despite protestations about building upon 'good practice' the Government-appointed SEAC (School Examination and Assessment Council) has the direct task of:

> —letting contracts for the production of standard assessment tasks . . .
> —advising [the Secretary of State] on the development, implementation and operation of the assessment system [my insertion in brackets]

These are but instances of the exceptionally rapid transference of power and control to the central state which has occurred over the last decade. In part this has been an attempt by central government to exert direct control of public expenditure. It has also been an assertive drive to impose national ('nationalist' being perhaps more appropriate) definitions, to destroy any pluralist power bases, and to muffle sources of alternative ideologies. As an outcome of this, teachers will increasingly be placed in the position of 'delivering' the national curriculum and 'executing' the assessment procedures, whilst LEAs will become 'administrative agencies' for centrally prescribed policies. As has been said about the Health Service White Paper (January 1989), it becomes a policy within which 'accountants and bureaucrats' are the key agents.

It would like to move towards concluding this review by claiming that there are quite different principles upon which it would be possible to base assessment procedures:

(1) They would attempt to combat rather than reinforce individualism. As such they would be based on principles of co-operation rather than competition.
(2) The purposes, procedures and criteria would be open to scrutiny and established collectively. This would include the opportunity for the assessed to respond to the results.
(3) Built into the procedures would be a recognition of the social and historical nature of knowledge. Also there would be a recognition that assessment procedures themselves are historically located and socially produced. As such they are products of power relations and link to wider social structures.
(4) Emphasis would be placed on the encouragement of positive results; clear indication of ways in which development might be made; working with the knowledge available in the group or community rather than focusing on abstract knowledge.
(5) Such procedures would be a component of a wider educational strategy to overcome repressive social divisions and not reinforce them.
(6) They would attempt to demystify expertise, professional knowledge and technicism rather than using these to legitimate results. This would also

mean taking into account the context, nature and purpose of the assessment situation.

Principles such as these would generate heated debate and controversy. They would overtly illustrate the manner in which educational issues slide inexorably into analysis of the kind of society/world we would wish to live in. Also it would be transparently clear that assessment constitutes only one element within any overall educational strategy. For exanmple, the teacher education course on which I work has a commitment to collaborative, active, pupil-centred learning, to State-maintained comprehensive schooling, and to the role of education in enhancing equalities. Pedagogy, politics, curriculum, ideology, justice and assessment—inevitably such terms slide and collide with each other and in so doing create what I will describe as 'democratic friction'. Is any society worth living in, or educational institution worth working in, ever going to be a 'frictionless' place?

To illustrate the relevance of this general point I will conclude with some questions which are quite different from those posed by the TGAT report:

—How could assessment be made a more social activity? Who discusses the issues that assessment raises: staff, parents, students, school governors; who else?

—Is it possible to involve students actively in the process of assessment? How? What kinds of conflicts, antagonisms, embarassment, friction and debate would this entail? Is it worth it?

—As educators, what are our motives for assessment, why do we do it? Do any of our reasons for undertaking it contradict one another, or other principles of education?

—Who is assessment for? Whose interests are served by it? Who gains, who loses? What happens to contradictions between interests?

Correspondence: Ian Hextall, University of London, Goldsmith' College, Department of Postgraduate Initial Teacher Education, New Cross, London SE14 6NW, United Kingdom.

NOTES

[1] See, for example, R. LEVITAS (Ed.) (1986) *The Ideology of The New Right* (Cambridge, Polity Press); and A. GAMBLE (1988) *The Free Economy and the Strong State* (London, Macmillan).

[2] A number of articles are beginning to attempt an overall analysis of the impact of Tory policies on education. See, for example, I. HEXTALL (1988) Education policy in England and Wales: the impact of the New Right, in: E. GUMBERT (Ed.) (1988) *The Politics of Educational Reform* (University of Georgia); G. WHITTY & I. MENTER (1989) Lessons of Thatcherism—education policy in England and Wales, 1979–1988, *Journal of Law and Society*, January; R. DALE (1988) Political change, educational change and the State in England 1944–1988, unpublished paper.

[3] V. WALKERDINE *et al.* (1988) *Girls and Mathematics: some lessons for the classroom* (ESRC); J. EGGLESTON *et al. Education for Some: the educational and vocational experiences of 15–18 year old members of minority ethnic groups*, esp. ch. 8; B. TIZARD & M. HUGHES (1984) *Young Children Learning* (London, Fontana). In their differing ways each of these articles refers to the implications of assessment procedures for gender, race and class divisions.

[4] Runnymede Trust and Radical Statistics Race Group (1980) *Britain's Black Population* (London, Heinemann), quoted in P. BROADFOOT (Ed.) (1984) *Selection, Certification and Control*, p. 259 (Lewes, Falmer Press).

[5] P. BOURDIEU (1976) The school as a conservative force: scholastic and cultural inequalities, in: R. DALE *et al.* (Eds) *Schooling and Capitalism, a Sociological Reader* (London, Routledge & Kegan Paul).

[6] Radical Statistics Education Group (1982) *Reading Between the Numbers* (BSSRS Publications). Quoted in P. BROADFOOT, op. cit.
[7] National Council for Civil Liberties (1989) *A Decade of Decline* (NCCL Publications).

The move towards integrating assessment more fully with teaching has been underway for some years in the United Kingdom. The educational arguments which underpin the move attend to issues of curriculum development and pupil learning, as well as more technical ones concerned with the validity and reliability of examination results. Thus, it has been claimed (and by many groups other than the currently much denigrated 'educational establishment') that contemporary schooling ought to focus more on developing new skills, capacities and understandings: for example, those of problem-solving, gathering and analysing data, and applying knowledge rather than simply remembering it. Teaching, and in turn validly assessing, such new skills and capacities demands the design of longer term and more relevant tasks for pupils to undertake (practical investigations, field trips, etc.) rather than traditional 'chalk and talk' followed by externally set and marked paper-and-pencil tests. Thus teachers are drawn into the assessment process formally as well as informally since they are in the best position to observe pupils at work over a period of time.

With regard to pupil learning, it is further argued that in addition to their curricular experience being rendered more meaningful by such changes in teaching, pupils will gain considerably by being offered regular 'formative feedback' on their strengths and weaknesses as identified by teachers when they carry out assessments *in situ*. Also some curriculum developments have adopted a 'graded objectives' approach to the pacing and sequencing of new curricula and claim that making shorter-term targets explicit to pupils increases their motivation to engage with the tasks.

For examiners the key issue, as indicated above, is being able to assess changing learning outcomes with validity, hence the move to assess work in the 'natural' setting of the laboratory or classroom, rather than the artificial setting of the examination hall. Additionally, however, the assessment of coursework and practical work can be said to contribute to the reliability of final grades by increasing the sample of assessed work on which judgements are made. The problem which examiners also have to face in this regard, however, is variability of school-based practice (in setting tasks as well as assessing them), hence the tendency (in GCSE for example) for the teacher role to be closely prescribed and the need for moderation to be undertaken.

This is the educational debate in which the arguments and recommendations of the Task Group on Assessment and Testing (TGAT) are grounded. The Task Group makes great play of the need for assessment to assist in the learning process and not merely be used to measure learning outcomes. It argues that for this to happen much more asessment information than simple test scores must be generated and, moreover, generated largely in the context of routine teacher–pupil interaction, so that pupils and parents can receive detailed 'attainment profiles' of pupils' strengths and weaknesses. As a consequence of such an argument, classroom teachers are given a central role in the recommended new

system, both in 'rating' their pupils' performance in the context of ordinary classroom work, and in conducting the 'standard assessment tasks' which the Task Group says ought to be devised and administered nationally to give the system a basis for directly comparing performance across schools and regions. In turn a local and regional consortium moderation procedure is recommended to monitor the final agreement of grades and to facilitate teacher development through discussion at local level. The Task Group's core recommendations are thus that assessment should be criterion-referenced, formative, assist pupil progression in learning, and be moderated so that the dual benefits of teacher involvement and national comparability can accrue.

However, the overall context in which such recommendations have been made has been dominated by the broader political debates which led to the changes in school governance included in the 1988 Education Reform Act, and by the much narrower view of assessment which the Department of Education and Science (DES) originally, and the Prime Minister subsequently, held and apparently still hold. Thus the DES in its original 'Consultation Document' stated that setting attainment targets was "a proven and essential way towards raising standards of achievement", that "at the heart of the assessment there will be nationally prescribed tests done by all pupils" and that "parents, governing bodies, employers and the local community should know what a school's assessment and examination results indicate about performance and how they compare with those of other schools . . ." (DES), 1987, pp. 10, 11, 14). Likewise the Prime Minister, in her leaked response to the first TGAT report, makes clear her preference for a simple system of targets and testing (*The Times Educational Supplement*, 18.3.88). Taken together with the provisions of the 1988 Act for open enrolment, local management of schools and schools opting out of direct local authority control altogether, the push towards straightforward testing carries many more messages about the production of simple performance indicators for the developing competitive market place in education than it does about raising standards through target-setting and curriculum specification—the ostensible reason for government action.

It is in the light of the damage that a narrow programme of testing could do to children's education, particularly in the primary sector, that the main TGAT report has been cautiously welcomed by many working in education. Yet while its recommendations are grounded in educational arguments, its choice of which arguments to highlight and which practices to subscribe to clearly owes a great deal to the prevailing political climate. Thus, although the report stresses the need to produce a broad range of assessment data, both to facilitate formative feedback to pupils, and to give greater confidence in overall results, it does so in parallel with its proposals for a particular model of graded cirricular progression such that teachers and pupils alike will be working towards targets over which they have no control and beyond which they will be extremely unlikely to stray. This betrays a highly instrumental and mechanistic view of learning and on which ignores broader definitions of formative assessment which would acknowledge the likelihood of unintended as well as intended learning outcomes being part of any educational encounter, and being worthy of identification and development in the context of more flexible curricular provision.

Likewise just as the parameters of legitimate pupil learning are tightly drawn by such a model so the main and supplementary TGAT reports construe the

professional development of teachers, as facilitated by involvement in school-based assessment and moderation, in a very restricted fashion. Far from being encouraged to analyse contemporary curricular problems for themselves, and think through the ways in which assessment might assist in the process of teaching and learning, teachers are simply taken to be operationally responsible for the quality of delivery of the new system and therefore regarded as having to be trained (or rather retrained) accordingly. Local Education Authorities (LEAs) and teachers are to be marshalled in a centre–periphery, 'cascade' approach to curriculum dissemination and in-service training with the time and resources necessary for this being met "by a shift in priorities so as to re-deploy the resources already being made available for INSET and related activities" (TGAT, 1988, p. 55). Now of course we can recognise that such phrases are the stuff of contemporary educational 'realpolitic'—it being hard enough to argue even a highly circumscribed educational case for a more sophisticated approach to assessment, let alone if it also costs more money—but then this is just the point of my argument, that for all the educational claims of TGAT, the parameters of what can count as a legitimate contribution to current debate are determined by other concerns.

All of which is not to suggest that the Task Group's reports have been self-consciously misleading—simply 'raiding' educational debate to legitimate political ends—but rather that a particularly selected group of individuals, operating in a particular set of circumstances, have attended to the more behaviourist end of the continuum of current assessment debate. Indeed it is still not absolutely clear what the Government's overriding concerns actually are. It seems to be attempting both to intervene in the curriculum in a peculiarly (given its rhetoric) corporatist fashion in order to prescribe 'standards' which will meet employers' concerns for a more skilled and competitive workforce in the future, while also laying the foundations of a highly competitive educational market place which, if it operates effectively, will of necessity respond to the much shorter-term concerns of current consumers (i.e. parents). In turn the TGAT reports have largely taken the corporatist stance of defining procedures and practices centrally, but have also attended to the issue of consumer choice by recommending the aggregation of individual, formative 'attainment profiles' into single, overall, subject-specific grades which will themselves be aggregated and published to indicate comparative class and school performance.

How the new system works out in practice remains to be seen. The potential contradictions within it are legion. For example, it is by no means clear that encouraging competition (between pupils and between schools) actually teaches people *how* to compete successfully. Indeed we know that, in the competitive system we have operated to date, the combination of school provision with pupil interest, ambitions, and most significantly, social circumstance, has led some pupils to learn how to succeed, but most how to fail. Perhaps this is why the government appears to have taken out an each-way bet on the corporatist and free-market approaches. At a more practical level, the TGAT proposals depend to a large extent on LEA structures and networks which would not survive large-scale opting out; the proposals depend even more on the goodwill and professionalism of teachers at a time when morale is low, resources are stretched and future employment and pay levels could depend on results achieved. In such circumstances 'teaching to the test' is inevitable, outright fraud a possibility (and one

which has surfaced under similar circumstances in the United States)—hardly the most appropriate way to develop a creative and flexible workforce for the future, let alone an informed and critical citizenry.

In the short term much will depend on the programmes of study which the National Curriculum Working Groups draw up, and the standard assessment tasks which the new examining consortia design. (In passing it is interesting to note that two members of the Task Group, Clare Bustall, Director of the National Foundation for Educational Research, and Tom Christie, a senior member of the Northern Examining Association, have been awarded £1.6m and £2.8m, respectively to develop the tasks they themselves were instrumental in recommending should be developed. We must never forget that examining pupils is itself a competitive business and investment in 'new technology' is always welcomed by the market leaders!). In the long term the publishing of results and the impact of parental choice on enrolment will have substantial implications for the organisation and quality of educational provision, as some schools shrink, others grow, and local authorities prevaricate over closures for fear of encouraging opting out.

Overall then, and without moving too far into the realms of speculation, it is clear that assessment has been alighted upon as an instrument which can affect both the process of schooling and its overall organisation. Control of the curriculum has been taken over by central government and the pattern of assessment now being developed will further strengthen that control by assigning specifically technical rather than professional roles to teachers; assessing pupils in particular ways, at particular times, on particular syllabuses will become a *de facto* part of a teacher's job specification. In turn the publication of results will impact on school management such that school organisation will change to maximise 'output' (the return of traditional streaming perhaps, or grade-related rather than age-related teaching groups). The irony of such a scenario is that it is unlikely to address in any real sense the key issues underlying the current crisis in comprehensive education—those of the quality of provision and accountability to the population at large. The government seems to believe that specification and measurement are the appropriate tools with which to impose quality and accountability. Yet not only is such an authoritarian strategy politically distasteful, it is also likely to be unsuccessful. It is too mechanistic in its basic conception, the more so when one considers the particular problems of specifying quality in as unpredictable an endeavour as education; and an endeavour, moreover, the defining characteristic of which is that it ought to be open in outcome, rather than closed. The alternative is to encourage professional development through far more intra-and inter-school discussion and to render accounts to parents through dialogue—both features of the Record of Achievement movement. In this regard the Task Group's recommendations for local moderation by discussion are the most positive feature of the proposed new system, although as noted previously, the brief for the consortium meetings is likely to be a very restricted one. To make the most of these meetings will require them being oriented towards far more than simply the agreement of grades. Similarly, talking to parents as openly as possible about the 'attainment profiles' of their own children ought to go some way towards creating a better understanding of what schools are doing and how they are doing it. Securing such benefits from the TGAT proposals will depend on substantial collaboration between teachers, parents, schools, local authorities and examining consortia. But of course the other

aspects of the Education Reform Act, outlined above, are in no sense designed to promote such collaboration.

Correspondence: Harry Torrance, Assessment and Evaluation Unit, School of Education, University of Southampton, Southampton SO9 5NH, United Kingdom.

Two aspects of government educational policy-making over the last few years now seem highly significant. The first is *pace*. Politicians and administrators compete to see who can move the fastest. The second is *orchestration*, and here I borrow the term from Michael Fullan. The build up, announcement and implementation of policy, even down to the accidents some suggest, is thought through with influential individuals or groups brought in or excluded in symphonic harmony or discord.

The tactics have been highly successful. The skills of political infighting are hardly the hallmark of the (genteel, down at heel?) educational establishment. New rules have been established, new agendas set. I am not averse to a faster pace, although it sits rather uncongenially alongside the liberal nostalgia for a collegial, participatory idea of change. It certainly offends professional opinion.

Against this scenario it is unsurprising that the TGAT report is not a consultative document. The task group, an appropriate new phrase in educational nomenclature, responded in just a few months to a Government request to establish a new national system of assessment and testing. The result is breathtaking. Only a few hedged bets. A simple and clear proposal which may prove one of the most influential educational documents of the post-war period. Why?

First, and one hardly dares say this in public, because it represented the first jolt to the momentum of official and unofficial Prime Ministerial advisors who had gained the high ground of educational policy makung during the build-up to the last election. The speed with which the report was produced and the orchestration of educational and public opinion outmanoeuvred those who would have had it strangled at birth (poor Professor Higginson's A level report, much longer in the making and let out at the wrong time was to suffer just such a fate).

The report's proposals do offer a comprehensible model for assessment which would give children, and their parents, a much greater stake in planning their curriculum. The growth of consumerism predates Thatcherism. Who would have thought that the Left's attacks on professionalism in the 1960s would have been taken up by a Tory Government in the 1980s?

Finally, the report is significant, not only for the opportunities it affords but also because of the grave problems that could develop if shortcomings remain unresolved.

Here we come up against one of the problems with debates about TGAT. The proposal for a 10-level system of criteria-based assessments was immediately latched upon by those fearful of crudely imposed alternatives. 'Don't rock the boat on TGAT for fear of something worse' echoed down the corridors of the

powerful and the sometimes unknowingly less than powerful. It was a dangerous message. TGAT needs critical scrutiny now.

Positive aspects of the proposals need exploiting. For me one of the most significant statements comes in one of the three supplementary reports (published as one volume) that followed a few months behind the publication of the main report. The task group state unequivocably:

> There has been some misunderstanding about the assessment of "ability", to which our main report may inadvertently have contributed by occasionally using that term. We had intended to confine our proposals to the assessment of "performance" or "attainment" and were not recommending any attempt to assess separately the problematic nature of underlying "ability". If "ability" were to be assessed, its meaning would have to be carefully defined; and the problem of defining it without making it merely the measure of a particular type of performance is hard to solve.

National assessment, therefore, is not about the ubiquitous '9' factor, discredited for a generation but alive and well in every staffroom in the country, but about what young people really have achieved. I think that is as significant a statement as Edward Boyle's famous 'acquiring intelligence' introduction to the Newson Report. TGAT could be about individual potential, about providing springboards to higher achievements rather than lids on known performance.

Next TGAT provides a structure for a curriculum map and a terminology which gives parents and their children access to the teaching plans. Schools have created a mystique around the curriculum, and an even bigger mystique around assessment. The proposals promise to break that down. It may represent a challenge to professional authority and autonomy but it provides a breath of liberating fresh air for millions who have been fobbed off with inadequate information about what they really can achieve. Most adults in this country believe they are only average or below average in potential. We do not want another generation with supressed expectations.

I suspect the real significance of both these points would be lost on many kitchen cabinet advisors. All due credit to TGAT members for the skill and ingenuity with which the concepts are delivered.

It is the problems, however, that need urgent attention. First a general point: critical scrutiny must extend beyond the assessment experts. There are too few of them! Sociologists (with some notable exceptions), no less than parents or children, have been baffled by assessment technicalities. Plenty of concern with outcomes, little attention to process. And yet, however sophisticated the terms or the statistics, underlying decisions of value are repeated at every stage in the process. If we accept, as I do, that parents and children have a right to some form of measurement of progress, then educationalists need to understand the methods and alternatives available. Six problems, therefore, with TGAT for you to mull over.

First, TGAT suggests that schools' performances should be checked against a sort of socio-economic vignette of the catchment area, if the results are to be interpreted fairly. Anyone who knows anything of the cut and thrust of parental choice of schools will know this to be an unrealistic, even nonsensical suggestion. If school performance is to be measured then the same technical statistical

procedures applied to student performance need translating into school terms. Any fair judgement of effort takes account of the level at which children enter the school are performing. Schools should be judged on progress made within the institution not on raw outcomes plus cameo descriptions describing leafy suburbia or run-down tenements.

Secondly, and here the technicalities come in, is the issue of aggregation. TGAT sets out to provide a series of statements about achievement within specific areas of knowledge, defined by TGAT as profile components. To arrive at a measure of performance in these profile components scores will be aggregated from a number of different attainment targets. The problem is that every time you aggregate you move one step away from recording what the child really can do. Higher education and public exams have happily done this for years. A 2:1 or 2:2 tells you how you rank with all the other students, it gives little information about what you know. The danger, therefore, is that the numbers assume too great an importance. This danger becomes particularly acute when the government, disregarding TGAT and working party advice, reduces the number of profile components *and* aggregates these scores into an overall level of subject performance. To illustrate the point a combined score on writing and oral communication would tell you little about the child's performance in either. I have a dread of entering an infant classroom in a few years time and being shown the table of Level 1 children, the table for Level 2 and the small group of Level 3 in the corner. The pressure of implementation is pushing in that direction.

Thirdly, there is the major inequality that can result from gender bias in assessment activities. My colleague Patricia Murphy (1989) has written extensively about this. The assessment of skills has to be contextualised and it is proving extremely difficult to find neutral stimuli. Place the assessment of a skill in the kitchen and the girls do better, in a garage and the boys leap ahead. Less obvious contexts prove equally problematic. The inclusion of a submission from the Equal Opportunities Commission as an appendix to the TGAT report hardly does justice to the significance of this issue.

Fourthly, evidence from APU has shown that the more practical an assessment activity the more likely the child is to display competence. Again TGAT fails to deal with the issue including a few rather haphazardly chosen examples. Practical activities require resources, and the government is already worried about cost implications, they also take time and here the teachers may prove an obstacle.

The report pays little attention to how realistic it is to expect teachers confidently to take on the new tasks. Proposals for in-service education in the supplementary report deal primarily with training outside the school. Excessive practical activities, although fairer to children, may prove difficult to implement even if resourcing is available. And, introducing a fifth problem, there is the issue of frequency of assessment. The more you assess the fairer the outcome. TGAT, however, rightfully mindful of the worry about over-assessment, adopts a minimalist approach. I am not clear how this difficulty is resolved. The different consortium currently developing Standard Assessment Tasks (SATs) may have an answer.

A sixth and final point is the extent to which parents and children will know what is happening. In TGAT there is the understandable concern to ensure that pupils experience success. Assessment, therefore, becomes integral to the teaching and childrn may not even be aware that formal assessment is taking place. I

think it unlikely that this is acceptable to a great many parents. There is, therefore, a conflict between the right to know and the attempt to integrate TGAT proposals into every teaching practice.

I suspect that all these problems will be starkly revealed in the development stages of the SATs. Three consortia are currently undertaking this work. The evaluation of the PILOT stage will be sigificant equaly so the first few years of national implementation. Sociologists of education played a powerful role in an older national system of assessment, the 11+. The present proposals are far more extensive, equally significant for life chances, hold out important opportunities but equally important dangers. A central issue, therefore, for debate and research.

Correspondence: Bob Moon, Centre for Curriculum and Teaching Studies, The Open University, Walton Hall, Milton Keynes MK7 6AA, United Kingdom.

REFERENCE

MURPHY, P. (1989) Gender and assessment, *Curriculum*, 9(3).