

Against the Stream

Harvey Goldstein and Richard Noss

Professor Harvey Goldstein and his colleague Richard Noss at the University of London Institute of Education examine the dubious concept of 'levels' and warn of its implications for schools.

The notion of 'levels' or 'stages' of learning has a long history — so long that it seems a natural framework for thinking about the organisation of the school curriculum. Ideas of ordering learning, from simple to complex, or from practical to abstract, seem useful for thinking about curriculum structure. Such ideas can be tested against experience, within different contexts and with different kinds of students. More problematic, however, is any assertion that a particular sequence of learning is necessary or optimum for everybody. Attempts to maintain such a stance have been unsuccessful, whether via an elaborate theoretical structure such as that of Piaget, or the more empirically based studies of graded assessments in mathematics and science (Hart, 1981). One might ask whether learning is like a mountain which has to be scaled, starting at the bottom and finishing at the top; or more like an exhibition which can be viewed in different orders by different people, with only a rather gentle pressure on everyone to walk around in the same direction.

Despite theoretical objections, the attraction of a neatly arranged set of levels through which all students should pass, recently seems to have had considerable attraction for those charged with the task of formulating educational policy. It clearly satisfies several requirements. First, it is simple and easy to describe. Secondly, and despite the research evidence, there are many educationalists who still share the view that the use of levels for an assessment scheme is educationally valid (Brown 1989). Thirdly, it provides a convenient administrative framework. Finally, and perhaps most importantly, it legitimates the segregation and streaming of children.

We shall pick up these issues in the following sections, but first we briefly outline the equally brief history of levels in the national curriculum.

TGAT and all that

Undoubtedly the key document in the development of the national curriculum and its assessment is the TGAT report (DES 1988). Set up shortly after the 1987 election, the TGAT committee was charged with fleshing out the existing policy commitment to testing at 7, 11, 14 and 16 years within a national curriculum. It was this report which established the 10 levels in each area of the curriculum to which all subsequent curriculum working parties have referred. While this report has been superseded by the many and varied working party reports that have succeeded it, it remains the only clear statement of the testing strategy which underpins the national curriculum.

The key recommendations concerning levels come in paragraphs 96-117 of the report. There we find the assertion that assessment 'gradings' should reflect a

child's 'progress' rather than just a child's ranking in comparison with other children. We are not told how anyone is to be prevented from using measures of progress in order to rank children. Indeed, almost everything in the report would encourage just such a use. By attaching a level assessment to each child, comparison with other children clearly is invited. Only by making such an assessment *private* between teacher and student can such comparisons be avoided. This is not what national assessment is about, however, with its plans to publish results for schools and the linking of levels to the GCSE grading system. All of which brings us to criterion referencing.

One of the most potent public notions in assessment during the 1980s is that of criterion referencing. Purveyed as an antidote to 'norm referenced' assessment which ranks students, criterion referencing is supposed to tell us what a student has 'mastered' or, in TGAT's words, 'understands, knows and can do'. The discussion of criterion referenced assessment is full of splendid sentiments about the importance of emphasising the 'positive' aspects of each student's achievements. Unfortunately, it is somewhat short on any critical examination of what this really means. In fact, while certain aspects of criterion referenced assessment might be useful, for example in defining reference domains, there is nothing which implies that these assessments have to be used differently from norm referenced assessments. Students can be ranked just as easily on the basis of a large series of 'mastery/non-mastery' judgements as on the basis of a small number of continuous test scores.

Thus criterion referencing will serve the purpose of public assessment just as well as any other system. Despite the rhetoric of the TGAT report and its successors, one must conclude that criterion referencing is used rather as a drunk uses a street lamp; for support rather than illumination!

Levels of attainment

Specifically, TGAT recommended that the 'average' child would move up one level every 2 years, starting from below level 1 up to a maximum of level 10. The higher levels would be equated with GCSE grades, so that there would be the possibility, indeed encouragement, to regard the level progression as part of the formal school-leaving certification system. Since everyone is supposed to start at roughly the same level in each area, the implication is that at any age there will be roughly the same distribution of levels within each area. Naturally, we cannot know whether this will be true until the system has been running for some time. Almost certainly it will not generally be true, and it is difficult to see how it could ever be made to happen

across all attainment targets and all curriculum subjects. Nevertheless, the definition by fiat does have important consequences for what happens in schools.

First, what is laid down as an average expectation is easily transmuted into a minimum requirement. It will often be the case that a student can satisfy some of the requirements for having achieved several levels, but not satisfy enough to be deemed to have arrived at any one of the levels. The danger is that such a student will be forced to concentrate on achieving the lowest level not yet attained in the necessary number of statements of attainment or attainment targets, before being allowed to move on. Hopefully, teachers will become aware of this danger. Whether the pressures will allow them to deal with it sensibly is another matter.

Secondly, if a teacher of seven-year-olds gets an average level of 2.1 in mathematics and an average of 1.7 in English, what is to be concluded? This might well be a common pattern, but will the system be flexible enough to discover this?

Thirdly, there is a danger that teachers will be encouraged to avoid the above situation by teaching to fulfil the TGAT prophecy. Namely to try to ensure that, on average, children do indeed progress at the same rate through every profile component or subject. That could lead to a severe imbalance.

Fourthly, a major omission in the TGAT report and subsequent discussions is any admission that there can be uncertainty or unreliability associated with any grading system. Variability between teachers, between the assessment tasks chosen for use and numerous other contingent factors mean that some uncertainty attaches to any statement of levels. A different task chosen, a different context in which a teacher assesses a child will often mean a different grading. That has to be remembered at the very least, even if the designers of the tests and the coordinators of the teachers' assessments give no guidance. Any numbers which become attached to children should therefore be viewed as approximations rather than precise statements.

Finally there is the whole issue of how schools and teachers will react to a 'high stakes' system in which a great deal hinges on maximising test scores or grades. The overt classification of children into levels, the linking of these to GCSE and the enormous demands on teacher time which is being asked, may well encourage streaming of children. This seems to us a very real danger, and we discuss it in more detail in the following section.

A return to streaming?

Given that so much depends on them, the proposals to publish school average scores or grades will encourage schools to use whatever devices they can to raise those averages. Already there is some evidence that schools might well perceive that the best way to maximise their overall test scores is to stream by achievement level. Such an approach has attractions for those who take seriously the notion of hierarchies of attainment. Indeed, if one is prepared to accept that there is an invariant sequence of learning attached to a particular topic, it does seem somewhat unlikely that pupils will progress at exactly the same rate and in exactly the

same way. Within such a philosophy, streaming *does* work: it produces hierarchies which reproduce themselves. Children in top streams *do* fare better than those in bottom streams — so streaming provides its own self-justification.

This kind of self-fulfilling prophesy is further compounded by the pedagogical approach which the national curriculum levels are likely to encourage. Since the nineteen-sixties, when 'mixed ability' teaching began to take root, there has been considerable confusion as to what the term might actually mean. While its early proponents argued for a style of teaching which took account of varying rates (and perhaps styles) of learning within a single class, the term 'mixed ability' has most often been used simply to mean the lumping together of children of a range of 'abilities' within a single class. In the extreme, this can involve the same chalk-and-talk methods which so singularly failed all but the top streams before the moves to mixed ability.

To a large extent, the commitment to mixed ability teaching came from below, as a response on the part of classroom teachers to the expectations raised by the introduction of comprehensive schooling in the fifties and sixties. Some of the best practice was and is in primary schools, where the number of pupils in any given year renders it unviable to have more than a single class. Significantly, the abolition of the 11-plus and the accompanying removal of the need to assess pupils competitively gave a boost to this process.

Despite some significant achievements at classroom level, the theory and practice of mixed-ability teaching is very varied, and sometimes confused. In primary classrooms it is not unusual to find children streamed into distinct groupings within a 'mixed ability' framework. In secondary schools it is common to use mixed ability groupings on entry as a mechanism subsequently to generate streams.

Thus there are two distinct aspects of mixed ability which have consistently been blurred. On the one hand there is mixed ability *grouping*: the placing of children of varying attainments within a single class. On the other there is mixed ability *teaching*: the explicit recognition that within a single class, it is possible and even desirable to accommodate different styles and rates of learning. Our concern here is with the latter, and the pressures which the imposition of the national curriculum will bring to bear to eradicate such approaches.

In secondary schools, the readiness of teachers to adopt mixed ability methods has been related to subject specialisms. For example, the teaching of foreign languages has been largely resistant to such methods, while there are plenty of examples of English (and, with regional variations, mathematics) teaching which adopt mixed ability approaches. There are a number of reasons for such diversity, and we do not consider them here. The point we wish to emphasise is that teachers' perceptions of the subject are more important than any intrinsic or psychological ordering of subject matter. The construction of invariant learning hierarchies in foreign language teaching, as much as their abolition in mathematics, reflects *pedagogical* priorities: it bears little relation to any particular property of the subject matter itself.

Whatever else the national curriculum does, it presents subject matter in a hierarchical and codified form, as sets of attainments which pupils are supposed to learn in a particular order. It is inevitable therefore, that this will affect the ways in which teachers are encouraged to think of their subjects, and of course, to teach it. The imposition of levels strongly reinforces the idea of progression through those levels. While in principle it might be argued that such progression is not tied to any particular style of teaching, the realities of the classroom are likely to suggest otherwise. The burden placed on teachers by the national curriculum is considerable (a fact even recently acknowledged by the government), and it is unrealistic to expect any but the most committed teachers to deliberately make their working lives even more burdensome by adopting the time-consuming practices associated with mixed-ability teaching. On the contrary, in subjects like mathematics, where the national curriculum involves a progression through 296 differentiated statements of attainment, the pressure will be to adopt not only organisational forms but also teaching styles which simplify the process of grading (see Dowling and Noss, 1990, for a critical review of the national curriculum in relation to mathematics).

Thus there is a pedagogical imperative associated with the national curriculum levels, to adopt a style of teaching which is at odds with mixed ability teaching. The reification of differences between pupils (objectified by attaching scores to performance) will inevitably lead to methods of teaching which reproduce the existence of those differences, and thus demonstrate the apparent unviability of teaching a wide range of attainment levels within a single class. Worse still, the emphasis on inter-individual differences in the form of public assessments may well lead to the blurring of intra-individual differences — and thus return us to the notion of ability as a quality of an individual which transcends subject boundaries. In this scenario, pupils will be branded as top-stream or bottom-stream at increasingly early ages.

Fundamental problems

If levels are to be given credibility some assessment system seems required. The oft repeated assertion is that the system allows parents, children, teachers and schools to 'know where they are'. Surely, it is argued, this is highly desirable?

The trouble is, 'knowing where you are' means that you will also know who is 'above' and 'below' you, with the problems we have already referred to. Yet, neither for schools nor for individual children, does the system tell anyone where they really are. As far as schools are concerned, the proposed system of reporting school averages manifestly will not enable schools to compare themselves validly with other schools (see Goldstein, 1990). Nor will the assigning of levels to children allow parents or others to know whether those particular children are receiving an adequate curriculum diet. Children differ in what they bring to school and in their ability to respond to teaching. An apparently 'poor' result does not imply that a child is being poorly taught or missing out on the curriculum. To make such a judgement is both unfair and seriously undermining of

teachers' professionalism. To be sure, the threat of such judgements may encourage teachers to devote themselves to raising test scores, but that is not to be confused with good teaching.

If there is a real concern to make sure that the components of a national curriculum is delivered, then assigning levels to students is not an efficient or an equitable method. Naturally, there is a serious requirement for students to understand their own learning progress and to take as much responsibility for their own learning as possible. A more appropriate vehicle for this however, would be a school based record of achievement scheme such as have now been tried in a number of places, with potential for expansion into primary schools. These would have no need of level assignments, but would be related to the curriculum which was taught, with important roles for different learning sequences, cross curricular work and extra curricular activities. Such schemes are essentially private, being a matter for discussion and compromise between the student, the teacher and the parents. As we have argued in detail elsewhere, the aims and implementation of private and public assessment schemes are fundamentally incompatible (see Noss, Goldstein and Hoyles, 1989).

To ensure that schools maintain overall curriculum standards, a more appropriate model is one where advisors, inspectors and the schools themselves cooperate to evaluate and improve their activities. This would ideally be done in a spirit of mutual concern for benefiting from everybody's experience, within an atmosphere of cooperation rather than one of wasteful, unnecessary and uninformative competition.

Such an approach is, however, entirely at odds with the intentions behind the national curriculum and its associated testing procedures. Indeed, we would argue that the national curriculum is primarily concerned with providing a common currency of test results with which to introduce the ethics and economics of the marketplace into the education system. In our view those who seek to mitigate the worst implementations of the system imposed by the 1988 Act, by striving to provide more 'humane' assessments, are misguided. Their motives often are laudable, but the system itself is so misdirected that such attempts at improvements are counter-productive. The most important task facing those who care about this country's education system is to find ways fundamentally to help teachers to swim against the anti-educational stream of the national curriculum, not to collaborate in its aims.

References

- Brown M. (1989) Graded Assessment and Learning Hierarchies in mathematics — an alternative view. *British Educational Research Journal*, Vol.15, 2, 121-128.
- DES (1988) *Task Group on Assessment and Testing* London: Dept. of Education and Science.
- Dowling P. and Noss R. (Eds). (1990). *Mathematics versus the National Curriculum* Basingstoke: Falmer Press.
- Goldstein H (1990) The fundamental assumptions of national assessment. In: Dowling P. and Noss R. (Eds) (1990). *Mathematics versus the National Curriculum*. Basingstoke: Falmer Press.
- Hart K. (Ed) (1981) *Children's Understanding of Mathematics 11-16*. London: J. Murray.
- Noss R., Goldstein H and Hoyles C. (1989) Graded Assessment and Learning Hierarchies in Mathematics. *British Education Research Journal*, 15, 2, 109-120, 1989